# SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: Hadlaczky, Gyula
Szalay, Aladar

(ii) TITLE OF THE INVENTION: ARTIFICIAL CHROMOSOMES, USES THEREOF AND METHODS FOR PREPARING ARTIFICIAL CHROMOSOMES

(iii) NUMBER OF SEQUENCES: 34

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Heller Ehrman White & McAuliffe
(B) STREET: 4250 Executive Square, 7th Floor
(C) CITY: La Jolla
(D) STATE: CA
(E) COUNTRY: USA
(F) ZIP: 92037

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: FastSEQ Version 1.5

09725094 . 112800

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE: 28-NOV-2000

(vi) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 08/835,682
(B) FILING DATE: 10-APR-1997
(C) CLASSIFICATION:

(vi) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 08/695,191
(B) FILING DATE: 07-AUG-1996
(C) CLASSIFICATION:

(vi) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 08/682,080
(B) FILING DATE: 15-JUL-1996
(C) CLASSIFICATION:

(vi) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 08/629,822
(B) FILING DATE: 10-APR-1996
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Seidman, Stephanie L
(B) REGISTRATION NUMBER: 33,779
(C) REFERENCE/DOCKET NUMBER: 24601-402G

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 858-450-8403
(B) TELEFAX: 858-587-5360
(C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1293 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:
(ix) FEATURE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
GAATTCATCA TTTTTCANGT CCTCAAGTGG ATGTTTCTCA TTTNCCATGA TTTTAAGTTT     60
TCTCGCCATA TTCCTGGTCC TACAGTGTGC ATTTCTCCAT TTTNCACGTT TTNCAGTGAT    120
TTCGTCATTT TCAAGTCCTC AAGTGGATGT TTCTCATTTN CCATGAATTT CAGTTTTCTN    180
GCCATATTCC ACGTCCTACA GNGGACATTT CTAAATTTNC CACCTTTTTC AGTTTTCCTC    240
GCCATATTTC ACGTCCTAAA ATGTGTATTT CTCGTTTNCC GTGATTTTCA GTTTTCTCGC    300
CAGATTCCAG GTCCTATAAT GTGCATTTCT CATTTNNCAC GTTTTTCAGT GATTTCGTCA    360
TTTTTTCAAG TCGGCAAGTG GATGTTTCTC ATTTNCCATG ATTTNCAGTT TTCTTGNAAT    420
ATTCCATGTC CTACAATGAT CATTTTTAAT TTTCCACCTT TTCATTTTTC CACGCCATAT    480
TTCATGTCCT AAAGTGTATA TTTCTCCTTT TCCGCGATTT TCAGTTTTCT CGCCATATTC    540
CAGGTCCTAC AGTGTGCATT CCTCATTTTT CACCTTTTTC ACTGATTTCG TCATTTTTCA    600
AGTCGTCAAC TGGATCTTTC TAATTTTCCA TGATTTTCAG TTATCTTGTC ATATTCCATG    660
TCCTACAGTG GACATTTCTA AATTTTCCAA CTTTTTCAAT TTTTCTCGAC ATATTTGACG    720
TGCTAAAGTG TGTATTTCTT ATTTTCCGTG ATTTTCAGTT TTCTCGCCAT ATTCCAGGTC    780
CTAATAGTGT GCATTTCTCA TTTTTCACGT TTTTCAGTGA TTTCGTCATT TTTTCCAGTT    840
GTCAAGGGGA TGTTTCTCAT TTTCCATGAG TGTCAGTTTT CTTGCTATAT TCCATGTCCT    900
ACAGTGACAT TTCTAAATAT TATACCTTTT TCAGTTTTTC TCACCATATT TCACGTCCTA    960
AAGTATATAT TTCTCATTTT CCCTGATTTT CAGTTTCCTT GCCATATTCC AGGTCCTACA   1020
GTGTGCATTT CTCATTTTTC ACGTTTTTCA GTAATTTCTT CATTTTTTAA GCCCTCAAAT   1080
GGATGTTTCT CATTTTCCAT GATTTTCAGT TTTCTTGCCA TATACCATGT CCTACAGTGG   1140
ACATTTCTAA ATTATCCACC TTTTTCAGTT TTTCATCGGC ACATTTCACG TCCTAAAGTG   1200
TGTATTTCTA ATTTTCAGTG ATTTTCAGTT TTCTCGCCAT ATTCCAGGAC CTACAGTGTG   1260
CATTTCTCAT TTTTCACGTT TTTCAGTGAA TTC                                1293
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1044 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:
(ix) FEATURE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
AGGCCTATGG TGAAAAAGGA AATATCTTCC CCTGAAAACT AGACAGAAGG ATTCTCAGAA     60
TCTTATTTGT GATGTGCGCC CCTCAACTAA CAGTGTTGAA GCTTTCTTTT GATAGAGCAG    120
TTTTGAAACA CTCTTTTTGT AAAATCTGCA AGAGGATATT TGGATAGCTT TGAGGATTTC    180
CGTTGGAAAC GGGATTGTCT TCATATAAAC CCTAGACAGA AGCATTCTCA GAAGCTTCAT    240
TGGGATGTTT CAGTTGAAGT CACAGTGTTG AACAGTCCCC TTTCATAGAG CAGGTTTGAA    300
ACACTCTTTT TTGTAGTATC TGGAAGTGGA CATTTGGAGC GATCTCAGGA CTGCGGTGAA    360
AAAGGAAATA TCTTCCAATA AAAGCTAGAT AGAGGCAATG TCAGAAACCT TTTTCATGAT    420
GTATCTACTC AGCTAACAGA GTTGAACCTT CCTTTGAGAG AGCAGTTTTG AAACACTCTT    480
TTTGTGGAAT CTGCAAGTGG ATATTTGTCT AGCTTTGAGG ATTTCGTTGG GAAACGGGAT    540
TACATATAAA AAGCAGACAG CAGCATTCCC AGAAACTTCT TTGTGATGTT TGCATTCAAG    600
TCACAGAGTT GAACATTCCC TTTCATAGAG CAGGTTTGAA ACACACTTTT TGATGTATCT    660
```

```
GGATGTGGAC ATTTGCAGCG CTTTCAGGCC TAAGGTGAAA AGGAAATATC TTCCCCTGAA     720
AACTAGACAG AAGCATTCTC AGAAACTTAT TTGTGATGTG CGCCCTCAAC TAACAGTGTT     780
GAAGCTTTCT TTTGATAGAG GCAGTTTTGA AACACTCTTT TGTGGAATCT GCAAGTGGAT     840
ATTTGTCTAG CTTTGAGGAT TTCTTTGGAA ACGGGATTAC ATATAAAAAG CAGACAGCAG     900
CATTCCCAGA ATCTTGTTTG TGATGTTTGC ATTCAAGTCA CAGAGTTGAA CATTCCCTTT     960
CAGAGAGCAG GTTTGAACAC TCTTTTTATA GTATCTGGAT GTGGACATTT GGAGCGCTTT    1020
CAGGGGGGAT CCTCTAGAAT TCCT                                          1044
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2492 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:
(ix) FEATURE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
CTGCAGCTGG GGGTCTCCAA TCAGGCAGGG GCCCCTTACT ACTCAGATGG GGTGGCCGAG      60
TAGGGGAAGG GGGTGCAGGC TGCATGAGTG GACACAGCTG TAGGACTACC TGGGGGCTGT     120
GGATCTATGG GGGTGGGGAG AAGCCCAGTG ACAGTGCCTA GAAGAGACAA GGTGGCCTGA     180
GAGGGTCTGA GGAACATAGA GCTGGCCATG TTGGGGCCAG GTCTCAAGCA GGAAGTGAGG     240
AATGGGACAG GCTTGAGGAT ACTCTACTCA GTAGCCAGGA TAGCAAGGAG GGCTTGGGGT     300
TGCTATCCTG GGGTTCAACC CCCCAGGTTG AAGGCCCTGG GGGAGATGGT CCCAGGACAT     360
ATTACAATGG ACACAGGAGG TTGGGACACC TGGAGTCACC AAACAAAACC ATGCCAAGAG     420
AGACCATGAG TAGGGGTGTC CAGTCCAGCC CTCTGACTGA GCTGCATTGT TCAAATCCAA     480
AGGGCCCCTG CTGCCACCTA GTGGCTGATG GCATCCACAT GACCCTGGGC CACACGCGTT     540
TAGGGTCTCT GTGAAGACCA AGATCCTTGT TACATTGAAC GACTCCTAAA TGAGCAGAGA     600
TTTCCACCTA TTCGAAACAA TCACATAAAA TCCATCCTGG AAAAAGCCTG GGGGATGGCA     660
CTAAGGCTAG GGATAGGGTG GGATGAAGAT TATAGTTACA GTAAGGGGTT TAGGGTTAGG     720
GATCAACGTT GGTTAGGAGT TAGGGATACA GTAGGGTACC GGTAGGGTTA GGGGTTAGGG     780
TTAGGGGTTA GGGTTAGGGT TAGGGTTAGG GTTAGGGTTA GGGGTTAGGG GTTAGGGTTA     840
GGGTTAGGTT TTGGGGTGGC GTATTTTGGT CTTATACGCT GTGTTCCACT GGCAATGAAA     900
AGAGTTCTTG TTTTTCCTTC AGCAATTTGT CATTTTTAAA AGAGTTTAGC AATTCTAACA     960
GATATAGACC AGCTGTGCTA TCTCATTGTG GTTTTCAATT GTAACCACAT TGTGGTTTCA    1020
ATGTGTTTAC TTGCCATCTG TAGATCTTCT TTGCGTGAGG TGTCTGTTCA GATGTGTGTG    1080
CATTTCTTGN NTTTNGGCTG TTTAACTTAT TGTTTAGTTT TAATAATTTT TTATATATTT    1140
GAAGACAAAT CTTTCTCAGA TGTGTATTTG CAAATATTTC TTCAATATGA GGCTTGCTTT    1200
TGTCTCTAAC AAGGTCTCTT CAGAGATAAC TTAAATATAA GAAATCCACA CTGTCACTTC    1260
TTTTGTGTAT ATCTACCTTT TGTGTCATTT GTTAAAATTC ATTACCAAAC CCAAAGGCAG    1320
ATAGCTTTTC TTCTATTGTT TCTTCTAGAA ATTTGTATAG TTTTGCATTT TTAGTGTAAG    1380
GATGATTTTG AGTGATTATT TGTGTAAGTT GTAAAGTTTT CGTCTATATC CATATCATTT    1440
CTTATGGTTT CCAATTAATC GTTCCCTCAC TATTTTTGGG AAAGACACAG GATAGTGGGC    1500
TTTGTTAGAG TAGATAGGTA GCTAGACATG AACAGGAGGG GGCCTCCTGG AAAAGGGAAA    1560
GTCTGGGAAG GCTCACCTGG AGGACCACCA AAAATTCACA TATTAGTAGC ATCTCTAGTG    1620
CTGGAGTGGA TGGGCACTTG TCAATTGTGG GTAGGAGGGA AAAGAGGTCC TATGCAGAAA    1680
GAAACTCCCT AGAACTCCTC TGAAGATGCC CCAATCATTC ACTCTGCAAT AAAAATGTCA    1740
GAATATTGCT AGCTACATGC TGATAAGGNN AAAGGGGACA TTCTTAAGTG AAACCTGGCA    1800
CCATAAGTAC AGATTAGGGC AGAGAAGGAC ATTCAAAAGA GGCAGGCGCA GTAGGTACAA    1860
ACGTGATCGC TGTCAGTGTG CCTGGGATGG CGGGAAGGAG GCTGGTGCCA GAGTGGATTC    1920
GTATTGATCA CCACACATAT ACCTCAACCA ACAGTGAGGA GGTCCCACAA GCCTAAGTGG    1980
GGCAAGTTGG GGAGCTAAGG CAGTAGCAGG AAAACCAGAC AAAGAAAACA GGTGGAGACT    2040
TGAGACAGAG GCAGGAATGT GAAGAAATCC AAAATAAAAT TCCCTGCACA GGACTCTTAG    2100
GCTGTTTAAT GCATCGCTCA GTCCCACTCC TCCCTATTTT TCTACAATAA ACTCTTTACA    2160
CTGTGTTTCT TTTCAATGAA GTTATCTGCC ATCTTTGTAT TGCCTCTTGG TGAAAATGTT    2220
TCTTCCAAGT TAAACAAGAA CTGGGACATC AGCTCTCCCC AGTAATAGCT CCGTTTCAGT    2280
```

TTGAATTTAC AGAACTGATG GGCTTAATAA CTGGCGCTCT GACTTTAGTG GTGCAGGAGG 2340
CCGTCACACC GGGACCAAGA GTGCCCTGCC TAGTCCCCAT CTGCCCGCAG GTGGCGGCTG 2400
CCTCGACACT GACAGCAATA GGGTCCGGCA GTGTCCCCAG CTGCCAGCAG GGGGCGTACG 2460
ACGACTACAC TGTGAGCAAG AGGGCCCTGC AG 2492

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:
(ix) FEATURE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GGGGAATTCA TTGGGATGTT TCAGTTGA 28

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:
(ix) FEATURE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CGAAAGTCCC CCCTAGGAGA TCTTAAGGA 29

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 47 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:
(ix) FEATURE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CCGCTTAATA CTCTGATGAG TCCGTGAGGA CGAAACGCTC TCGCACC

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:
(ix) FEATURE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CGATTTAAAT TAATTAAGCC CGGGC 25

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:
(ix) FEATURE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

TAAATTTAAT TAATTCGGGC CCGTCGA 27

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 69 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(D) OTHER INFORMATION IL-2 signal sequence

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

ATG TAC AGG ATG CAA CTC CTG TCT TGC ATT GCA CTA AGT CTT GCA CTT 48
Met Tyr Arg Met Gln Leu Leu Ser Cys Ile Ala Leu Ser Leu Ala Leu

GTC ACA AAC AGT GCA CCT ACT 69
Val Thr Asn Ser Ala Pro Thr

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 945 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence
(B) LOCATION: 1...942
(D) OTHER INFORMATION: Renilla Reinformis Luciferase

(x) PUBLICATION INFORMATION:

PATENT NO.: 5,418,155

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

```
AGC TTA AAG ATG ACT TCG AAA GTT TAT GAT CCA GAA CAA AGG AAA CGG      48
Ser Leu Lys Met Thr Ser Lys Val Tyr Asp Pro Glu Gln Arg Lys Arg
 1               5                  10                  15

ATG ATA ACT GGT CCG CAG TGG TGG GCC AGA TGT AAA CAA ATG AAT GTT      96
Met Ile Thr Gly Pro Gln Trp Trp Ala Arg Cys Lys Gln Met Asn Val
            20                  25                  30

CTT GAT TCA TTT ATT AAT TAT TAT GAT TCA GAA AAA CAT GCA GAA AAT     144
Leu Asp Ser Phe Ile Asn Tyr Tyr Asp Ser Glu Lys His Ala Glu Asn
        35                  40                  45

GCT GTT ATT TTT TTA CAT GGT AAC GCG GCC TCT TCT TAT TTA TGG CGA     192
Ala Val Ile Phe Leu His Gly Asn Ala Ala Ser Ser Tyr Leu Trp Arg
    50                  55                  60

CAT GTT GTG CCA CAT ATT GAG CCA GTA GCG CGG TGT ATT ATA CCA GAT     240
His Val Val Pro His Ile Glu Pro Val Ala Arg Cys Ile Ile Pro Asp
65                  70                  75                  80

CTT ATT GGT ATG GGC AAA TCA GGC AAA TCT GGT AAT GGT TCT TAT AGG     288
Leu Ile Gly Met Gly Lys Ser Gly Lys Ser Gly Asn Gly Ser Tyr Arg
                85                  90                  95

TTA CTT GAT CAT TAC AAA TAT CTT ACT GCA TGG TTG AAC TTC TTA ATT     336
Leu Leu Asp His Tyr Lys Tyr Leu Thr Ala Trp Leu Asn Phe Leu Ile
            100                 105                 110

TAC CAA AGA AGA TCA TTT TTT GTC GGC CAT GAT TGG GGT GCT TGT TTG     384
Tyr Gln Arg Arg Ser Phe Phe Val Gly His Asp Trp Gly Ala Cys Leu
        115                 120                 125

GCA TTT CAT TAT AGC TAT GAG CAT CAA GAT AAG ATC AAA GCA ATA GTT     432
Ala Phe His Tyr Ser Tyr Glu His Gln Asp Lys Ile Lys Ala Ile Val
    130                 135                 140

CAC GCT GAA AGT GTA GTA GAT GTG ATT GAA TCA TGG GAT GAA TGG CCT     480
His Ala Glu Ser Val Val Asp Val Ile Glu Ser Trp Asp Glu Trp Pro
145                 150                 155                 160

GAT ATT GAA GAA GAT ATT GCG TTG ATC AAA TCT GAA GAA GGA GAA AAA     528
Asp Ile Glu Glu Asp Ile Ala Leu Ile Lys Ser Glu Glu Gly Glu Lys
                165                 170                 175

ATG GTT TTG GAG AAT AAC TTC TTC GTG GAA ACC ATG TTG CCA TCA AAA     576
Met Val Leu Glu Asn Asn Phe Phe Val Glu Thr Met Leu Pro Ser Lys
            180                 185                 190

ATC ATG AGA AAG TTA GAA CCA GAA GAA TTT GCA GCA TAT CTT GAA CCA     624
```

```
Ile Met Arg Lys Leu Glu Pro Glu Glu Phe Ala Ala Tyr Leu Glu Pro
        195                 200                 205

TTC AAA GAG AAA GGT GAA GTT CGT CGT CCA ACA TTA TCA TGG CCT CGT       672
Phe Lys Glu Lys Gly Glu Val Arg Arg Pro Thr Leu Ser Trp Pro Arg
    210                 215                 220

GAA ATC CCG TTA GTA AAA GGT GGT AAA CCT GAC GTT GTA CAA ATT GTT       720
Glu Ile Pro Leu Val Lys Gly Gly Lys Pro Asp Val Val Gln Ile Val
225                 230                 235                 240

AGG AAT TAT AAT GCT TAT CTA CGT GCA AGT GAT GAT TTA CCA AAA ATG       768
Arg Asn Tyr Asn Ala Tyr Leu Arg Ala Ser Asp Asp Leu Pro Lys Met
                245                 250                 255

TTT ATT GAA TCG GAT CCA GGA TTC TTT TCC AAT GCT ATT GTT GAA GGC       816
Phe Ile Glu Ser Asp Pro Gly Phe Phe Ser Asn Ala Ile Val Glu Gly
            260                 265                 270

GCC AAG AAG TTT CCT AAT ACT GAA TTT GTC AAA GTA AAA GGT CTT CAT       864
Ala Lys Lys Phe Pro Asn Thr Glu Phe Val Lys Val Lys Gly Leu His
        275                 280                 285

TTT TCG CAA GAA GAT GCA CCT GAT GAA ATG GGA AAA TAT ATC AAA TCG       912
Phe Ser Gln Glu Asp Ala Pro Asp Glu Met Gly Lys Tyr Ile Lys Ser
    290                 295                 300

TTC GTT GAG CGA GTT CTC AAA AAT GAA CAA TAA                           945
Phe Val Glu Arg Val Leu Lys Asn Glu Gln
305                 310
```

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:
(ix) FEATURE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

```
TTTGAATTC A TGTACAGGAT GCAACTCCTG                                     30
```

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:
(ix) FEATURE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

```
TTTGAATTCA GTAGGTGCAC TGTTTGTCAC                                        30
```

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1434 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

```
CCTCCACGCA CGTTGTGATA TGTAGATGAT AATCATTATC AGAGCAGCGT TGGGGGATAA      60
TGTCGACATT TCCACTCCCA ATGACGGTGA TGTATAATGC TCAAGTATTC TCCTGCTTTT     120
TTACCACTAA CTAGGAACTG GGTTTGGCCT TAATTCAGAC AGCCTTGGCT CTGTCTGGAC     180
AGGTCCAGAC GACTGACACC ATTAACACTT TGTCAGCCTC AGTGACTACA GTCATAGATG     240
AACAGGCCTC AGCTAATGTC AAGATACAGA GAGGTCTCAT GCTGGTTAAT CAACTCATAG     300
ATCTTGTCCA GATACAACTA GATGTATTAT GACAAATAAC TCAGCAGGGA TGTGAACAAA     360
AGTTTCCGGG ATTGTGTGTT ATTTCCATTC AGTATGTTAA ATTTACTAGG ACAGCTAATT     420
TGTCAAAAAG TCTTTTTCAG TATATGTTAC AGAATTGGAT GGCTGAATTT GAACAGATCC     480
TTCGGGAATT GAGACTTCAG GTCAACTCCA CGCGCTTGGA CCTGTCGCTG ACCAAAGGAT     540
TACCCAATTG GATCTCCTCA GCATTTTCTT TCTTTAAAAA ATGGGTGGGA TTAATATTAT     600
TTGGAGATAC ACTTTGCTGT GGATTAGTGT TGCTTCTTTG ATTGGTCTGT AAGCTTAAGG     660
CCCAAACTAG GAGAGACAAG GTGGTTATTG CCCAGGCGCT TGCAGGACTA GAACATGGAG     720
CTTCCCCTGA TATATGGTTA TCTATGCTTA GGCAATAGGT CGCTGGCCAC TCAGCTCTTA     780
TATCCCACGA GGCTAGTCTC ATTGTACGGG ATAGAGTGAG TGTGCTTCAG CAGCCCGAGA     840
GAGTTGCAAG GCTAAGCACT GCAATGGAAA GGCTCTGCGG CATATATGTG CCTATTCTAG     900
GGGGACATGT CATCTTTCAT GAAGGTTCAG TGTCCTAGTT CCCTTCCCCC AGGCAAAACG     960
ACACGGGAGC AGGTCAGGGT TGCTCTGGGT AAAAGCCTGT GAGCCTGGGA GCTAATCCTG    1020
TACATGGCTC CTTTACCTAC ACACTGGGGA TTTGACCTCT ATCTCCACTC TCATTAATAT    1080
GGGTGGCCTA TTTGCTCTTA TTAAAAGGAA AGGGGGAGAT GTTGGGAGCC GCGCCCACAT    1140
TCGCCGTTAC AAGATGGCGC TGACAGCTGT GTTCTAAGTG GTAAACAAAT AATCTGCGCA    1200
TGTGCCGAGG GTGGTTCTTC ACTCCATGTG CTCTGCCTTC CCCGTGACGT CAACTCGGCC    1260
GATGGGCTGC AGCCAATCAG GGAGTGACAC GTCCTAGGCG AAGGAGAATT CTCCTTAATA    1320
GGGACGGGGT TTCGTTCTCT CTCTCTCTCT TGCTTCTCTC TCTTGCTTTT TCGCTCTCTT    1380
GCTTCCCGTA AAGTGATAAT GATTATCATC TACATATCAC AACGTGCGTG GAGG          1434
```

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1400 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

```
CCTCCACGCA CGTTGTGATA TGTAGATGAT AATCATTATC AGAGCAGCGT TGGGGGATAA      60
TGTCGACATT TCCACTCCCA ATGACGGTGA TGTATAATGC TCAAGTATTC TCCTGCTTTT     120
TTACCACTAA CTAGGAACTG GGTTTGGCCT TAATTCAGAC AGCCTTGGCT CTGTCTGGAC     180
AGGTCCAGAT ACAACTAGAT GTATTATGAC AAATAACTCA GCAGGGATGT GAACAAAAGT     240
```

```
TTCCGGGATT GCGTGTTATT TCCATCCAGT ATGTTAAATT TACTAGGGCA GCTAATTTGT     300
CAAAAAGTCT TTTCCAGTAT ATGTTACAGA ATTGGATGGC TGAATTTGAA CAGATCCTTC     360
GGGAATTGAG ACTTCAGGTC AACTCCACGC GCTTGGACCT GTCCCTGACC AAAGGATTAC     420
CCAATTGGAT CTCCTCAGCA TTTTCTTTCT TTAAAAAATG GGTGGGATTA ATATTATTTG     480
GAGATACACT TTGCTGTGGA TTAGTGTTGC TTCTTTGATT GGTCTGTAAG CTTAAGGCCC     540
AAACTAGGAG AGACAAGGTG GTTATTGCCC AGGCGCTTGC AGGACTAGAA CATGGAGCTT     600
CCCCTGATAT ATCTATGCTT AGGCAATAGG TCGCTGGCCA CTCAGCTCTT ATATCCCATG     660
AGGCTAGTCT CATTGCACGG GATAGAGTGA GTGTGCTTCA GCAGCCCGAG AGAGTTGCAC     720
GGCTAAGCAC TGCAATGGAA AGGCTCTGCG GCATATATGA GCCTATTCTA GGGAGACATG     780
TCATCTTTCA AGAAGGTTGA GTGTCCAAGT GTCCTTCCTC CAGGCAAAAC GACACGGGAG     840
CAGGTCAGGG TTGCTCTGGG TAAAAGCCTG TGAGCCTAAG AGCTAATCCT GTACATGGCT     900
CCTTTACCTA CACACTGGGG ATTTGACCTC TATCTCCACT CTCATTAATA TGGGTGGCCT     960
ATTTGCTCTT ATTAAAAGGA AAGGGGGAGA TGTTGGGAGC CGCGCCCACA TTCGCCGTTA    1020
CAAGATGGCG CTGACAGCTG TGTTCTAAGT GGTAAACAAA TAATCTGCGC ATGCGCCGAG    1080
GGTGGTTCTT CACTCCATGT GCTCTGCCTT CCCCGTGACG TCAACTCGGC CGATGGGCTG    1140
CAGTCAATCA GGGAGTGACA CGTCCTAGGC GAAGGAAAAT TCTCCTTAAT AGGGACGGGG    1200
TTTCGTTTTC TCTCTCTCTT GCTTCGCTCT CTCTTGCTTC TTGCTCTCTT TTCCTGAAGA    1260
TGTAAGAATA AAGCTTTGCC GCAGAAGATT CTGGTCTGTG GTGTTCTTCC TGGCCGGTCG    1320
TGAGAACGCG TCTAATAACA ATTGGTGCCG AAACCCGGGT GATAATGATT ATCATCTACA    1380
TATCACAACG TGCGTGGAGG                                                1400
```

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1369 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

```
CCTCCACGCA CGTTGTGATA TGTAGATGAT AATCATTATC ACTTTACGGG TCCTTTCACT      60
ACAACTGCCA CGAGGCCCCG TGCTCTGGTA ATAGATCTTT GCTGAAAAGG CACACACATG     120
ACACATTACT CAAGGTGGGC TCATCTGAGC TGCAGATTCA GCTTAATATG AATCTTGCCA     180
ATTGTGTGAA ATCATAAATC TTCAAAGTGA CACTCATTGC CAGACACAGG TGCCCACCTT     240
TGGCATAATA AACAAACACA AATTATCTAT TATATAAAGG GTGTTAGAAG ATGCTTTAGA     300
ATACAAATAA ATCATGGTAG ATAACAGTAA GTTGAGAGCT TAAATTTAAT AAAGTGATAT     360
ACCTAATAAA AATTAAATTA AGAAGGTGTG AATATACTAC AGTAGGTAAA TTATTTCATT     420
AATTTATTTT CTTTCTTAAT CCTTTATAAT GTTTTCTGCT ATTGTCAATT GCACATCCAT     480
ATGTTCAATT CTTCACTGTA ATGAAGAAAT GTAGTAAATA TACTTTCCGA ACAAGTTGTA     540
TCAAATATGT TACACTTGAT TCCGTGTGTT ACTTATCATT TTATTATTAT ATTGATTGCA     600
TTCCTTCGTT ACTTGATATT ATTACAAGGT ACATATTTAT TCTCTCAGAT CTTCATTATA     660
CTCTAACCAT TTTATAACAT ACTTTATTTA TTCATTTCTT ATGTGTGCTG TGAGGCACAA     720
ATGCCAGAGA GAACTTGAGC AGATAAGAGG ACAAATTGCA AGAGTCAGTT ACCTCCTGCT     780
GTTCCTTGGA AACTCAGGAT CAAATTCAGG TTGTCAGGCT TGGCAGCATG CACTTTTTAC     840
CAGTGCCTCC ATCTTGCTAG CCCTGAACAT CAAGCTTTGC AGACAGACAG GCTACACTAA     900
GTGAACTGGT CATTCACAGC ATGCATGGTG ATTTATTGTT ACTTTCTATT CCATGCCTTT     960
ACTATTTCTA CTAGGTGCTA GCTAGTACTG TATTTCGAGA TAGAAGTTAC TGAAAGAAAA    1020
TTACATTGTT TTCTATAGAT CCTTGATACT CTTTCAGCAG ATATAGAGTT TTAATCAGGT    1080
CCTAGACCCT TTCTTCACTC TTATTAAATA CTAAGTACAA ATTAAGTTTA TCCAAAACAG    1140
TACGGATGTT GATTTTGTGC AGTTCTACTA TGATAATAGT CTAGCTTCAT AAATCTGACA    1200
CACTTATTGG GAATGTTTTT GTTAATAAAA GATTCAGGTG TTACTCTAGG TCAAGAGAAT    1260
ATTAAACATC AGTCCCAAAT TACAAACTTC AATAAAAGAT TTGACTCTCC AGTGGTGGCA    1320
ATATAAAGTG ATAATGATTA TCATCTACAT ATCACAACGT GCGTGGAGG                1369
```

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22118 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

```
GAATTCCCCT ATCCCTAATC CAGATTGGTG GAATAACTTG GTATAGATGT TTGTGCATTA     60
AAAACCCTGT AGGATCTTCA CTCTAGGTCA CTGTTCAGCA CTGGAACCTG AATTGTGGCC    120
CTGAGTGATA GGTCCTGGGA CATATGCAGT TCTGCACAGA CAGACAGACA GACAGACAGA    180
CAGACAGACA GACAGACGTT ACAAACAAAC ACGTTGAGCC GTGTGCCAAC ACACACACAA    240
ACACCACTCT GGCCATAATT ATTGAGGACG TTGATTTATT ATTCTGTGTT TGTGAGTCTG    300
TCTGTCTGTC TGTCTGTCTG TCTGTCTGTC TATCAAACCA AAAGAAACCA AACAATTATG    360
CCTGCCTGCC TGCCTGCCTG CCTACACAGA GAAATGATTT CTTCAATCAA TCTAAAACGA    420
CCTCCTAAGT TTGCCTTTTT TCTCTTTCTT TATCTTTTTC TTTTTTCTTT TCTTCTTCCT    480
TCCTTCCTTC CTTCCTTCCT TCCTTCCTTT CTTTCTTTCT TTCTTTCTTT CTTACTTTCT    540
TTCTTTCCTT CTTACATTTA TTCTTTTCAT ACATAGTTTC TTAGTGTAAG CATCCCTGAC    600
TGTCTTGAAG ACACTTTGTA GGCCTCAATC CTGTAAGAGC CTTCCTCTGC TTTTCAAATG    660
CTGGCATGAA TGTTGTACCT CACTATGACC AGCTTAGTCT TCAAGTCTGA GTTACTGGAA    720
AGGAGTTCCA AGAAGACTGG TTATATTTTT CATTTATTAT TGCATTTTAA TTAAAATTTA    780
ATTTCACCAA AAGAATTTAG ACTGACCAAT TCAGAGTCTG CCGTTTAAAA GCATAAGGAA    840
AAAGTAGGAG AAAAACGTGA GGCTGTCTGT GGATGGTCGA GGCTGCTTTA GGGAGCCTCG    900
TCACCATTCT GCACTTGCAA ACCGGGCCAC TAGAACCCGG TGAAGGGAGA AACCAAAGCG    960
ACCTGGAAAC AATAGGTCAC ATGAAGGCCA GCCACCTCCA TCTTGTTGTG CGGGAGTTCA   1020
GTTAGCAGAC AAGATGGCTG CCATGCACAT GTTGTCTTTC AGCTTGGTGA GGTCAAAGTA   1080
CAACCGAGTC ACAGAACAAG GAAGTATACA CAGTGAGTTC CAGGTCAGCC AGAGTTTACA   1140
CAGAGAAACC ACATCTTGAA AAAAACAAAA AAATAAATTA AATAAATATA ATTTAAAAAT   1200
TTAAAAATAG CCGGGAGTGA TGGCGCATGT CTTTAATCCC AGCTCTCTTC AGGCAGAGAT   1260
GGGAGGATTT CTGAGTTTGA GGCCAGCCTG GTCTGCAAAG TGAGTTCCAG GACAGTCAGG   1320
GCTATACAGA GAAACCCTGT CTTGAAAACT AAACTAAATT AAACTAAACT AAACTAAAAA   1380
AATATAAAAT AAAAATTTTA AAGAATTTTA AAAAACTACA GAAATCAAAC ATAAGCCCAC   1440
GAGATGGCAA GTAACTGCAA TCATAGCAGA AATATTATAC ACACACACAC ACACAGACTC   1500
TGTCATAAAA TCCAATGTGC CTTCATGATG ATCAAATTTC GATAGTCAGT AATACTAGAA   1560
GAATCATATG TCTGAAAATA AAAGCCAGAA CCTTTTCTGC TTTTGTTTTC TTTTGCCCCA   1620
AGATAGGGTT TCTCTCAGTG TATCCCTGGC ATCCCTGCCT GGAACTTCCT TTGTAGGTTT   1680
GGTAGCCTCA AACTCAGAGA GGTCCTCTCT GCCTGCCTGC CTGCCTGCCT GCCTGCCTGC   1740
CTGCCTGCCT GCCTGCCTCA CTTCTTCTGC CACCCACACA ACCGAGTCGA ACCTAGGATC   1800
TTTATTTCTT TCTCTTTCTC TCTTCTTTCT TTCTTTCTTT CTTTCTTTCT TTCTTTCTTT   1860
CTTTCTTTCT TTCTTATTCA ATTAGTTTTC AATGTAAGTG TGTGTTTGTG CTCTATCTGC   1920
TGCCTATAGG CCTGCTTGCC AGGAGAGGGC AACAGAACCT AGGAGAAACC ACCATGCAGC   1980
TCCTGAGAAT AAGTGAAAAA ACAACAAAAA AAGGAAATTC TAATCACATA GAATGTAGAT   2040
ATATGCCGAG GCTGTCAGAG TGCTTTTTAA GGCTTAGTGT AAGTAATGAA AATTGTTGTG   2100
TGTCTTTTAT CCAAACACAG AAGAGAGGTG GCTCGGCCTG CATGTCTGTT GTCTGCATGT   2160
AGACCAGGCT GGCCTTGAAC ACATTAATCT GTCTGCCTCT GCTTCCCTAA TGCTGCGATT   2220
AAAGGCATGT GCCACCACTG CCCGGACTGA TTTCTTCTTT TTTTTTTTTT TGGAAAATAC   2280
CTTTCTTTCT TTTTCTCTCT CTCTTTCTTC CTTCCTTCCT TTCTTTCTAT TCTTTTTTTC   2340
TTTCTTTTTT CTTTTTTTTT TTTTTTTTAA AATTTGCCTA AGGTTAAAGG TGTGCTCCAC   2400
AATTGCCTCA GCTCTGCTCT AATTCTCTTT AAAAAAAAAC AAACAAAAAA AAAACCAAAA   2460
CAGTATGTAT GTATGTATAT TTAGAAGAAA TACTAATCCA TTAATAACTC TTTTTTCCTA   2520
AAATTCATGT CATTCTTGTT CCACAAAGTG AGTTCCAGGA CTTACCAGAG AAACCCTGTG   2580
TTCAAATTTC TGTGTTCAAG GTCACCCTGG CTTACAAAGT GAGTTCCAAG TCCGATAGGG   2640
CTACACAGAA AAACCATATC TCAGAAAAAA AAAAAGTTCC AAACACACAC ACACACACAC   2700
ACACACACAC ACACACACAC ACACACACAC ACACACACAG CGCGCCGCGG CGATGAGGGG   2760
AAGTCGTGCC TAAAATAAAT ATTTTTCTGG CCAAAGTGAA AGCAAATCAC TATGAAGAGG   2820
TACTCCTAGA AAAAATAAAT ACAAACGGGC TTTTTAATCA TTCCAGCACT GTTTTAATTT   2880
AACTCTGAAT TTAGTCTTGG AAAAGGGGGC GGGTGTGGGT GAGTGAGGGC GAGCGAGCAG   2940
ACGGGCGGGC GGGCGGGTGA GTGGCCGGCG GCGGTGGCAG CGAGCACCAG AAAACAACAA   3000
ACCCCAAGCG GTAGAGTGTT TTAAAAATGA GACCTAAATG TGGTGGAACG GAGGTCGCCG   3060
CCACCCTCCT CTTCCACTGC TTAGATGCTC CCTTCCCCTT ACTGTGCTCC CTTCCCCTAA   3120
CTGTGCCTAA CTGTGCCTGT TCCCTCACCC CGCTGATTCG CCAGCGACGT ACTTTGACTT   3180
```

```
CAAGAACGAT TTTGCCTGTT TTCACCGCTC CCTGTCATAC TTTCGTTTTT GGGTGCCCGA   3240
GTCTAGCCCG TTCGCTATGT TCGGGCGGGA CGATGGGGAC CGTTTGTGCC ACTCGGGAGA   3300
AGTGGTGGGT GGGTACGCTG CTCCGTCGTG CGTGCGTGAG TGCCGGAACC TGAGCTCGGG   3360
AGACCCTCCG GAGAGACAGA ATGAGTGAGT GAATGTGGCG GCGCGTGACG GATCTGTATT   3420
GGTTTGTATG GTTGATCGAG ACCATTGTCG GGCGACACCT AGTGGTGACA AGTTTCGGGA   3480
ACGCTCCAGG CCTCTCAGGT TGGTGACACA GGAGAGGGAA GTGCCTGTGG TGAGGCGACC   3540
AGGGTGACAG GAGGCCGGGC AAGCAGGCGG GAGCGTCTCG GAGATGGTGT CGTGTTTAAG   3600
GACGGTCTCT AACAAGGAGG TCGTACAGGG AGATGGCCAA AGCAGACCGA GTTGCTGTAC   3660
GCCCTTTTGG GAAAAATGCT AGGGTTGGTG GCAACGTTAC TAGGTCGACC AGAAGGCTTA   3720
AGTCCTACCC CCCCCCCCCT TTTTTTTTTT TTTCCTCCAG AAGCCCTCTC TTGTCCCCGT   3780
CACCGGGGGC ACCGTACATC TGAGGCCGAG AGGACGCGAT GGGCCCGGCT TCCAAGCCGG   3840
TGTGGCTCGG CCAGCTGGCG CTTCGGGTCT TTTTTTTTTT TTTTTTTTTT TTTTCCTCCA   3900
GAAGCCTTGT CTGTCGCTGT CACCGGGGGC GCTGTACTTC TGAGGCCGAG AGGACGCGAT   3960
GGGCCCCGGC TTCCAAGCCG GTGTGGCTCG GCCAGCTGGA GCTTCGGGTC TTTTTTTTTT   4020
TTTTTTTTTT TTTTTTTCTC CAGAAGCCTT GTCTGTCGCT GTCACCGGGG GCGCTGTACT   4080
TCTGAGGCCG AGAGGACGCG ATGGGTCGGC TTCCAAGCCG ATGTGGCGGG GCCAGCTGGA   4140
GCTTCGGGTT TTTTTTTTTC CTCCAGAAGC CCTCTCTTGT CCCCGTCACC GGGGGCGCTG   4200
TACTTCTGAG GCCGAGAGGA CGTGATGGGC CCGGGTTCCA GGCGGATGTC GCCCGGTCAG   4260
CTGGAGCTTT GGATCTTTTT TTTTTTTTTT CCTCCAGAAG CCCTCTCTTG TCCCCGTCAC   4320
CGGGGGCACC TTACATCTGA GGGCGAGAGG ACGTGATGGG TCCGGCTTCC AAGCCGATGT   4380
GGCGGGGCCA GCTGGAGCTT CGGGTTTTTT TTTTTTCCTC CAGAAGCCCT CTCTTGTCCC   4440
CGTCACCGGG GGCGCTGTAC TTCTGAGGCC GAGAGGACGT GATGGGCCCG GGTTCCAGGC   4500
GGATGTCGCC CGGTCAGCTG GAGCTTTGGA TCATTTTTTT TTTTCCCTCC AGAAGCCCTC   4560
TCTTGTCCCC GTCACCGGGG GCACCGTACA TCTGAGGCCG AGAGGACACG ATGGGCCTGT   4620
CTTCCAAGCC GATGTGGCCC GGCCAGCTGG AGCTTCGGGT CTTTTTTTTT TTTTTTCCTC   4680
CAGAAGCCTT GTCTGTCGCT GTCACCCGGG GCGCTGTACT TCTGAGGCCG AGAGGACGCG   4740
ATGGGCCCGG CTTCCAAGCC GGTGTGGCTC GGCCAGCTGG AGCTTCGGGT CTTTTTTTTT   4800
TTTTTTTTTT TTCCTCCAGA AACCTTGTCT GTCGCTGTCA CCCGGGGCGC TTGTACTTCT   4860
GATGCCGAGA GGACGCGATG GGCCCGTCTT CCAGGCCGAT GTGGCCCGGT CAGCTGGAGC   4920
TTTGGATCTT TTTTTTTTTT TTTTCCTCCA GAAGCCCTCT CTTGTCCCCG TCACCGGGGG   4980
CACCTTACAT CTGAGGCCTA GAGGACACGA TGGGCCCGGG TTCCAGGCCG ATGTGGCCCG   5040
GTCAGCTGGA GCTTTGGATC TTTTTTTTTT TTTTCTTCCA GAAGCCCTCT TGTCCCCGTC   5100
ACCGGTGGCA CTGTACATCT GAGGCGGAGA GGACATTATG GGCCCGGCTT CCAATCCGAT   5160
GTGGCCCGGT CAGCTGGAGC TTTGGATCTT ATTTTTTTTT TAATTTTTTC TTCCAGAAGC   5220
CCTCTTGTCC CTGTCACCGG TGGCACGGTA CATCTGAGGC CGAGAGGACA TTATGGGCCC   5280
GGCTTCCAGG CCGATGTGGC CCGGTCAGCT GGAGCTTTGG ATCTTTTTTT TTTTTTTTCT   5340
TTTTTCCTCC AGAAGCCCTC TCTGTCCCTG TCACCGGGGG CCCTGTACGT CTGAGGCCGA   5400
GGGAAAGCTA TGGGCGCGGT TTTCTTTCAT TGACCTGTCG GTCTTATCAG TTCTCCGGGT   5460
TGTCAGGGTC GACCAGTTGT TCCTTTGAGG TCCGGTTCTT TTCGTTATGG GGTCATTTTT   5520
GGGCCACCTC CCCAGGTATG ACTTCCAGGC GTCGTTGCTC GCCTGTCACT TTCCTCCCTG   5580
TCTCTTTTAT GCTTGTGATC TTTTCTATCT GTTCCTATTG GACCTGGAGA TAGGTACTGA   5640
CACGCTGTCC TTTCCCTATT AACACTAAAG GACACTATAA AGAGACCCTT TCGATTTAAG   5700
GCTGTTTTGC TTGTCCAGCC TATTCTTTTT ACTGGCTTGG GTCTGTCGCG GTGCCTGAAG   5760
CTGTCCCCGA GCCACGCTTC CTGCTTTCCC GGGCTTGCTG CTTGCGTGTG CTTGCTGTGG   5820
GCAGCTTGTG ACAACTGGGC GCTGTGACTT TGCTGCGTGT CAGACGTTTT TCCCGATTTC   5880
CCCGAGGTGT CGTTGTCACA CCTGTCCCGG TTGGAATGGT GGAGCCAGCT GTGGTTGAGG   5940
GCCACCTTAT TTCGGCTCAC TTTTTTTTTT TTTTTTTCTC TTGGAGTCCC GAACCTCCGC   6000
TCTTTTCTCT TCCCGGTCTT TCTTCCACAT GCCTCCCGAG TGCATTTCTT TTTGTTTTTT   6060
TTCTTTTTTT TTTTTTTTTT TTGGGGAGGT GGAGAGTCCC GAGTACTTCA CTCCTGTCTG   6120
TGGTGTCCAA GTGTTCATGC CACGTGCCTC CCGAGTGCAC TTTTTTTTGT GGCAGTCGCT   6180
CGTTGTGTTC TCTTGTTCTG TGTCTGCCCG TATCAGTAAC TGTCTTGCCC CGCGTGTAAG   6240
ACATTCCTAT CTCGCTTGTT TCTCCCGATT GCGCGTCGTT GCTCACTCTT AGATCGATGT   6300
GGTGCTCCGG AGTTCTCTTC GGGCCAGGGC CAAGCCGCGC CAGGCGAGGG ACGGACATTC   6360
ATGGCGAATG GCGGCCGCTC TTCTCGTTCT GCCAGCGGGC CCTCGTCTCT CCACCCCATC   6420
CGTCTGCCGG TGGTGTGTGG AAGGCAGGGG TGCGGCTCTC CGGCCCGACG CTGCCCCGCG   6480
CGCACTTTTC TCAGTGGTTC GCGTGGTCCT TGTGGATGTG TGAGGCGCCC GGTTGTGCCC   6540
TCACGTGTTT CACTTTGGTC GTGTCTCGCT TGACCATGTT CCCAGAGTCG GTGGATGTGG   6600
CCGGTGGCGT TGCATACCCT TCCCGTCTGG TGTGTGCACG CGCTGTTTCT TGTAAGCGTC   6660
GAGGTGCTCC TGGAGCGTTC CAGGTTTGTC TCCTAGGTGC CTGCTTCTGA GCTGGTGGTG   6720
GCGCTCCCCA TTCCCTGGTG TGCCTCCGGT GCTCCGTCTG GCTGTGTGCC TTCCCGTTTG   6780
TGTCTGAGAA GCCCGTGAGA GGGGGGTCGA GGAGAGAAGG AGGGGCAAGA CCCCCCTTCT   6840
TCGTCGGGTG AGGCGCCCAC CCCGCGACTA GTACGCCTGT GCGTAGGGCT GGTGCTGAGC   6900
GGTCGCGGCT GGGGTTGGAA AGTTTCTCGA GAGACTCATT GCTTTCCCGT GGGGAGCTTT   6960
GAGAGGCCTG GCTTTCGGGG GGGACCGGTT GCAGGGTCTC CCCTGTCCGC GGATGCTCAG   7020
AATGCCCTTG GAAGAGAACC TTCCTGTTGC CGCAGACCCC CCCGCGCGGT CGCCCGCGTG   7080
```

```
TTGGTCTTCT GGTTTCCCTG TGTGCTCGTC GCATGCATCC TCTCTCGGTG GCCGGGGCTC   7140
GTCGGGGTTT TGGGTCCGTC CCGCCCTCAG TGAGAAAGTT TCCTTCTCTA GCTATCTTCC   7200
GGAAAGGGTG CGGGCTTCTT ACGGTCTCGA GGGGTCTCTC CCGAATGGTC CCCTGGAGGG   7260
CTCGCCCCCT GACCGCCTCC CGCGCGCGCA GCGTTTGCTC TCTCGTCTAC CGCGGCCCGC   7320
GGCCTCCCCG CTCCGAGTTC GGGGAGGGAT CACGCGGGGC AGAGCCTGTC TGTCGTCCTG   7380
CCGTTGCTGC GGAGCATGTG GCTCGGCTTG TGTGGTTGGT GGCTGGGGAG AGGGCTCCGT   7440
GCACACCCCC GCGTGCGCGT ACTTTCCTCC CCTCCTGAGG GCCGCCGTGC GGACGGGGTG   7500
TGGGTAGGCG ACGGTGGGCT CCCGGGTCCC CACCCGTCTT CCCGTGCCTC ACCCGTGCCT   7560
TCCGTCGCGT GCGTCCCTCT CGCTCGCGTC CACGACTTTG GCCGCTCCCG CGACGGCGGC   7620
CTGCGCCGCG CGTGGTGCGT GCTGTGTGCT TCTCGGGCTG TGTGGTTGTG TCGCCTCGCC   7680
CCCCCCTTCC CGCGGCAGCG TTCCCACGGC TGGCGAAATC GCGGGAGTCC TCCTTCCCCT   7740
CCTCGGGGTC GAGAGGGTCC GTGTCTGGCG TTGATTGATC TCGCTCTCGG GGACGGGACC   7800
GTTCTGTGGG AGAACGGCTG TTGGCCGCGT CCGGCGCGAC GTCGGACGTG GGGACCCACT   7860
GCCGCTCGGG GGTCTTCGTC GGTAGGCATC GGTGTGTCGG CATCGGTCTC TCTCTCGTGT   7920
CGGTGTCGCC TCCTCGGGCT CCCGGGGGGC CGTCGTGTTT CGGGTCGGCT CGGCGCTGCA   7980
GGTGTGGTGG GACTGCTCAG GGGAGTGGTG CAGTGTGATT CCCGCCGGTT TTGCCTCGCG   8040
TGCCCTGACC GGTCCGACGC CCGAGCGGTC TCTCGGTCCC TTGTGAGGAC CCCCTTCCGG   8100
GAGGGGCCCG TTTCGGCCGC CCTTGCCGTC GTCGCCGGCC CTCGTTCTGC TGTGTCGTTC   8160
CCCCCTCCCC GCTCGCCGCA GCCGGTCTTT TTTCCTCTCT CCCCCCCTCT CCTCTGACTG   8220
ACCCGTGGCC GTGCTGTCGG ACCCCCCGCA TGGGGGCGGC CGGGCACGTA CGCGTCCGGG   8280
CGGTCACCGG GGTCTTGGGG GGGGGCCGAG GGGTAAGAAA GTCGGCTCGG CGGGCGGGAG   8340
GAGCTGTGGT TTGGAGGGCG TCCCGGCCCC GCGGCCGTGG CGGTGTCTTG CGCGGTCTTG   8400
GAGAGGGCTG CGTGCGAGGG GAAAAGGTTG CCCCGCGAGG GCAAAGGGAA AGAGGCTAGC   8460
AGTGGTCATT GTCCCGACGG TGTGGTGGTC TGTTGGCCGA GGTGCGTCTG GGGGGCTCGT   8520
CCGGCCCTGT CGTCCGTCGG GAAGGCGCGT GTTGGGGCCT GCCGGAGTGC CGAGGTGGGT   8580
ACCCTGGCGG TGGGATTAAC CCCGCGCGCG TGTCCCGGTG TGGCGGTGGG GGCTCCGGTC   8640
GATGTCTACC TCCCTCTCCC CGAGGTCTCA GGCCTTCTCC GCGCGGGCTC TCGGCCCTCC   8700
CCTCGTTCCT CCCTCTCGCG GGGTTCAAGT CGCTCGTCGA CCTCCCCTCC TCCGTCCTTC   8760
CATCTCTCGC GCAATGGCGC CGCCCGAGTT CACGGTGGGT TCGTCCTCCG CCTCCGCTTC   8820
TCGCCGGGGG CTGGCCGCTG TCCGGTCTCT CCTGCCCGAC CCCCGTTGGC GTGGTCTTCT   8880
CTCGCCGGCT TCGCGGACTC CTGGCTTCGC CCGGAGGGTC AGGGGGCTTC CCGGTTCCCC   8940
GACGTTGCGC CTCGCTGCTG TGTGCTTGGG GGGGGCCCGC TGCGGCCTCC GCCCGCCCGT   9000
GAGCCCCTGC CGCACCCGCC GGTGTGCGGT TTCGCGCCGC GGTCAGTTGG GCCCTGGCGT   9060
TGTGTCGCGT CGGGAGCGTG TCCGCCTCGC GGCGGCTAGA CGCGGGTGTC GCCGGGCTCC   9120
GACGGGTGGC CTATCCAGGG CTCGCCCCCG CCGACCCCCG CCTGCCCGTC CCGGTGGTGG   9180
TCGTTGGTGT GGGGAGTGAA TGGTGCTACC GGTCATTCCC TCCCGCGTGG TTTGACTGTC   9240
TCGCCGGTGT CGCGCTTCTC TTTCCGCCAA CCCCCACGCC AACCCACCAC CCTGCTCTCC   9300
CGGCCCGGTG CGGTCGACGT TCCGGCTCTC CCGATGCCGA GGGGTTCGGG ATTTGTGCCG   9360
GGGACGGAGG GGAGAGCGGG TAAGAGAGGT GTCGGAGAGC TGTCCCGGGG CGACGCTCGG   9420
GTTGGCTTTG CCGCGTGCGT GTGCTCGCGG ACGGGTTTTG TCGGACCCCG ACGGGGTCGG   9480
TCCGGCCGCA TGCACTCTCC CGTTCCGCGC GAGCGCCCGC CCGGCTCACC CCCGGTTTGT   9540
CCTCCCGCGA GGCTCTCCGC CGCCGCCGCC TCCTCCTCCT CTCTCGCGCT CTCTGTCCCG   9600
CCTGGTCCTG TCCCACCCCC GACGCTCCGC TCGCGCTTCC TTACCTGGTT GATCCTGCCA   9660
GGTAGCATAT GCTTGTCTCA AAGATTAAGC CATGCATGTC TAAGTACGCA CGGCCGGTAC   9720
AGTGAAACTG CGAATGGCTC ATTAAATCAG TTATGGTTCC TTTGGTCGCT CGCTCCTCTC   9780
CTACTTGGAT AACTGTGGTA ATTCTAGAGC TAATACATGC CGACGGGCGC TGACCCCCCT   9840
TCCCGGGGGG GGATGCGTGC ATTTATCAGA TCAAAACCAA CCCGGTGAGC TCCCTCCCGG   9900
CTCCGGCCGG GGGTCGGGCG CCGGCGGCTT GGTGACTCTA GATAACCTCG GGCCGATCGC   9960
ACGCCCCCCG TGGCGGCGAC GACCCATTCG AACGTCTGCC CTATCAACTT TCGATGGTAG  10020
TCGCCGTGCC TACCATGGTG ACCACGGGTG ACGGGGAATC AGGGTTCGAT TCCGGAGAGG  10080
GAGCCTGAGA AACGGCTACC ACATCCAAGG AAGGCAGCAG GCGCGCAAAT TACCCACTCC  10140
CGACCCGGGG AGGTAGTGAC GAAAAATAAC AATACAGGAC TCTTTCGAGG CCCTGTAATT  10200
GGAATGAGTC CACTTTAAAT CCTTTAACGA GGATCCATTG GAGGGCAAGT CTGGTGCCAG  10260
CAGCCGCGGT AATTCCAGCT CCAATAGCGT ATATTAAAGT TGCTGCAGTT AAAAAGCTCG  10320
TAGTTGGATC TTGGGAGCGG GCGGGCGGTC CGCCGCGAGG CGAGTCACCG CCCGTCCCCG  10380
CCCCTTGCCT CTCGGCGCCC CCTCGATGCT CTTAGCTGAG TGTCCCGCGG GGCCCGAAGC  10440
GTTTACTTTG AAAAAATTAG AGTGTTCAAA GCAGGCCCGA GCCGCCTGGA TACCGCAGCT  10500
AGGAATAATG GAATAGGACC GCGGTTCTAT TTTGTTGGTT TTCGGAACTG AGGCCATGAT  10560
TAAGAGGGAC GGCCGGGGGC ATTCGTATTG CGCCGCTAGA GGTGAAATTC TTGGACCGGC  10620
GCAAGACGGA CCAGAGCGAA AGCATTTGCC AAGAATGTTT TCATTAATCA AGAACGAAAG  10680
TCGGAGGTTC GAAGACGATC AGATACCGTC GTAGTTCCGA CCATAAACGA TGCCGACTGG  10740
CGATGCGGCG GCGTTATTCC CATGACCCGC CGGGCAGCTT CCGGGAAACC AAAGTCTTTG  10800
GGTTCCGGGG GGAGTATGGT TGCAAAGCTG AAACTTAAAG GAATTGACGG AAGGGCACCA  10860
CCAGGAGTGG GCCTGCGGCT TAATTTGACT CAACACGGGA AACCTCACCC GGCCCGGACA  10920
CGGACAGGAT TGACAGATTG ATAGCTCTTT CTCGATTCCG TGGGTGGTGG TGCATGGCCG  10980
```

```
TTCTTAGTTG GTGGAGCGAT TTGTCTGGTT AATTCCGATA ACGAACGAGA CTCTGGCATG  11040
CTAACTAGTT ACGCGACCCC CGAGCGGTCG GCGTCCCCCA ACTTCTTAGA GGGACAAGTG  11100
GCGTTCAGCC ACCCGAGATT GAGCAATAAC AGGTCTGTGA TGCCCTTAGA TGTCCGGGGC  11160
TGCACGCGCG CTACACTGAC TGGCTCAGCG TGTGCCTACC CTGCGCCGGC AGGCGCGGGT  11220
AACCCGTTGA ACCCCATTCG TGATGGGGAT CGGGGATTGC AATTATTCCC CATGAACGAG  11280
GAATTCCCAG TAAGTGCGGG TCATAAGCTT GCGTTGATTA AGTCCCTGCC CTTTGTACAC  11340
ACCGCCCGTC GCTACTACCG ATTGGATGGT TTAGTGAGGC CCTCGGATCG GCCCCGCCGG  11400
GGTCGGCCCA CGGCCCTGGC GGAGCGCTGA GAAGACGGTC GAACTTGACT ATCTAGAGGA  11460
AGTAAAAGTC GTAACAAGGT TTCCGTAGGT GAACCTGCGG AAGGATCATT AAACGGAGA  11520
CTGTGGAGGA GCGGCGGCGT GGCCCGCTCT CCCCGTCTTG TGTGTGTCCT CGCCGGGAGG  11580
CGCGTGCGTC CCGGGTCCCG TCGCCCGCGT GTGGAGCGAG GTGTCTGGAG TGAGGTGAGA  11640
GAAGGGGTGG GTGGGGTCGG TCTGGGTCCG TCTGGGACCG CCTCCGATTT CCCCTCCCCC  11700
TCCCCTCTCC CTCGTCCGGC TCTGACCTCG CCACCCTACC GCGGCGGCGG CTGCTCGCGG  11760
GCGTCTTGCC TCTTTCCCGT CCGGCTCTTC CGTGTCTACG AGGGGCGGTA CGTCGTTACG  11820
GGTTTTTGAC CCGTCCCGGG GGCGTTCGGT CGTCGGGGCG CGCGCTTTGC TCTCCCGGCA  11880
CCCATCCCCG CCGCGGCTCT GGCTTTTCTA CGTTGGCTGG GGCGGTTGTC GCGTGTGGGG  11940
GGATGTGAGT GTCGCGTGTG GGCTCGCCCG TCCCGATGCC ACGCTTTTCT GGCCTCGCGT  12000
GTCCTCCCCG CTCCTGTCCC GGGTACCTAG CTGTCGCGTT CCGGCGCGGA GGTTTAAGGA  12060
CCCCGGGGGG GTCGCCCTGC CGCCCCCAGG GTCGGGGGGC GGTGGGGCCC GTAGGGAAGT  12120
CGGTCGTTCG GGCGGCTCTC CCTCAGACTC CATGACCCTC CTCCCCCCGC TGCCGCCGTT  12180
CCCGAGGCGG CGGTCGTGTG GGGGGGTGGA TGTCTGGAGC CCCCTCGGGC GCCGTGGGGG  12240
CCCGACCCGC GCCGCCGGCT TGCCCGATTT CCGCGGGTCG GTCCTGTCGG TGCCGGTCGT  12300
GGGTTCCCGT GTCGTTCCCG TGTTTTTCCG CTCCCGACCC TTTTTTTTTC CTCCCCCCCA  12360
CACGTGTCTC GTTTCGTTCC TGCTGGCCGG CCTGAGGCTA CCCCTCGGTC CATCTGTTCT  12420
CCTCTCTCTC CGGGGAGAGG AGGGCGGTGG TCGTTGGGGG ACTGTGCCGT CGTCAGCACC  12480
CGTGAGTTCG CTCACACCCG AAATACCGAT ACGACTCTTA GCGGTGGATC ACTCGGCTCG  12540
TGCGTCGATG AAGAACGCAG CTAGCTGCGA GAATTAATGT GAATTGCAGG ACACATTGAT  12600
CATCGACACT TCGAACGCAC TTGCGGCCCC GGGTTCCTCC CGGGGCTACG CCTGTCTGAG  12660
CGTCGGTTGA CGATCAATCG CGTCACCCGC TGCGGTGGGT GCTGCGCGGC TGGGAGTTTG  12720
CTCGCAGGGC CAACCCCCCA ACCCGGGTCG GGCCCTCCGT CTCCCGAAGT TCAGACGTGT  12780
GGGCGGTTGT CGGTGTGGCG CGCGCGCCCG CGTCGCGGAG CCTGGTCTCC CCCGCGCATC  12840
CGCGCTCGCG GCTTCTTCCC GCTCCGCCGT TCCCGCCCTC GCCCGTGCAC CCCGGTCCTG  12900
GCCTCGCGTC GGCGCCTCCC GGACCGCTGC CTCACCAGTC TTTCTCGGTC CCGTGCCCCG  12960
TGGGAACCCA CCGCGCCCCC GTGGCGCCCG GGGGTGGGCG CGTCCGCATC TGCTCTGGTC  13020
GAGGTTGGCG GTTGAGGGTG TGCGTGCGCC GAGGTGGTGG TCGGTCCCCT GCGGCCGCGG  13080
GGTTGTCGGG GTGGCGGTCG ACGAGGGCCG GTCGGTCGCC TGCGGTGGTT GTCTGTGTGT  13140
GTTTGGGTCT TGCGCTGGGG GAGGCGGGGT CGACCGCTCG CGGGGTTGGC GCGGTCGCCC  13200
GGCGCCGCGC ACCCTCCGGC TTGTGTGGAG GGAGAGCGAG GGCGAGAACG GAGAGAGGTG  13260
GTATCCCCGG TGGCGTTGCG AGGGAGGGTT TGGCGTCCCG CGTCCGTCCG TCCCTCCCTC  13320
CCTCGGTGGG CGCCTTCGCG CCGCACGCGG CCGCTAGGGG CGGTCGGGGC CCGTGGCCCC  13380
CGTGGCTCTT CTTCGTCTCC GCTTCTCCTT CACCCGGGCG GTACCCGCTC CGGCGCCGGC  13440
CCGCGGGACG CCGCGGCGTC CGTGCGCCGA TGCGAGTCAC CCCCGGGTGT TGCGAGTTCG  13500
GGGAGGGAGA GGGCCTCGCT GACCCGTTGC GTCCCGGCTT CCCTGGGGGG GACCCGGCGT  13560
CTGTGGGCTG TGCGTCCCGG GGGTTGCGTG TGAGTAAGAT CCTCCACCCC CGCCGCCCTC  13620
CCCTCCCGCC GGCCTCTCGG GGACCCCCTG AGACGGTTCG CCGGCTCGTC CTCCCGTGCC  13680
GCCGGGTGCC GTCTCTTTCC CGCCCGCCTC CTCGCTCTCT TCTTCCCGCG GCTGGGCGCG  13740
TGTCCCCCCT TTCTGACCGC GACCTCAGAT CAGACGTGGC GACCCGCTGA ATTTAAGCAT  13800
ATTAGTCAGC GGAGGAAAAG AAACTAACCA GGATTCCCTC AGTAACGGCG AGTGAACAGG  13860
GAAGAGCCCA GCGCCGAATC CCCGCCGCGC GTCGCGGCGT GGGAAATGTG GCGTACGGAA  13920
GACCCACTCC CCGGCGCCGC TCGTGGGGGG CCCAAGTCCT TCTGATCGAG GCCCAGCCCG  13980
TGGACGGTGT GAGGCCGGTA GCGGCCCCGG CGCGCCGGGC TCGGGTCTTC CCGGAGTCGG  14040
GTTGCTTGGG AATGCAGCCC AAAGCGGGTG GTAAACTCCA TCTAAGGCTA AATACCGGCA  14100
CGAGACCGAT AGTCAACAAG TACCGTAAGG GAAAGTTGAA AAGAACTTTG AAGAGAGAGT  14160
TCAAGAGGGC GTGAAACCGT TAAGAGGTAA ACGGGTGGGG TCCGCGCAGT CCGCCCGGAG  14220
GATTCAACCC GGCGGCGCGC GTCCGGCCGT GCCCGGTGGT CCCGGCGGAT CTTTCCCGCT  14280
CCCCGTTCCT CCCGACCCCT CCACCCGCGC GTCGTTCCCC TCTTCCTCCC CGCGTCCGGC  14340
GCCTCCGGCG GCGGGCGCGG GGGGTGGTGT GGTGGTGGCG CGCGGGCGGG GCCGGGGGTG  14400
GGGTCGGCGG GGGACCGCCC CCGGCCGGCG ACCGGCCGCC GCCGGGCGCA CTTCCACCGT  14460
GGCGGTGCGC CGCGACCGGC TCCGGGACGG CCGGGAAGGC CCGGTGGGGA AGGTGGCTCG  14520
GGGGGGGCGG CGCGTCTCAG GGCGCGCCGA ACCACCTCAC CCCGAGTGTT ACAGCCCTCC  14580
GGCCGCGCTT TCGCCGAATC CCGGGGCCGA GGAAGCCAGA TACCCGTCGC CGCGCTCTCC  14640
CTCTCCCCCC GTCCGCCTCC CGGGCGGGCG TGGGGGTGGG GGCCGGGCCG CCCCTCCCAC  14700
GGCGCGACCG CTCTCCCACC CCCCTCCGTC GCCTCTCTCG GGGCCCGGTG GGGGGCGGGG  14760
CGGACTGTCC CCAGTGCGCC CCGGGCGTCG TCGCGCCGTC GGGTCCCGGG GGGACCGTCG  14820
GTCACGCGTC TCCCGACGAA GCCGAGCGCA CGGGGTCGGC GGCGATGTCG GCTACCCACC  14880
```

```
CGACCCGTCT TGAAACACGG ACCAAGGAGT CTAACGCGTG CGCGAGTCAG GGGCTCGTCC  14940
GAAAGCCGCC GTGGCGCAAT GAAGGTGAAG GGCCCCGCCC GGGGGCCCGA GGTGGGATCC  15000
CGAGGCCTCT CCAGTCCGCC GAGGGCGCAC CACCGGCCCG TCTCGCCCGC CGCGCCGGGG  15060
AGGTGGAGCA CGAGCGTACG CGTTAGGACC CGAAAGATGG TGAACTATGC TTGGGCAGGG  15120
CGAAGCCAGA GGAAACTCTG GTGGAGGTCC GTAGCGGTCC TGACGTGCAA ATCGGTCGTC  15180
CGACCTGGGT ATAGGGGCGA AAGACTAATC GAACCATCTA GTAGCTGGTT CCCTCCGAAG  15240
TTTCCCTCAG GATAGCTGGC GCTCTCGCTC CCGACGTACG CAGTTTTATC CGGTAAAGCG  15300
AATGATTAGA GGTCTTGGGG CCGAAACGAT CTCAACCTAT TCTCAAACTT TAAATGGGTA  15360
AGAAGCCCGG CTCGCTGGCG TGGAGCCGGG CGTGGAATGC GAGTGCCTAG TGGGCCACTT  15420
TTGGTAAGCA GAACTGGCGC TGCGGGATGA ACCGAACGCC GGGTTAAGGC GCCCGATGCC  15480
GACGCTCATC AGACCCCAGA AAAGGTGTTG GTTGATATAG ACAGCAGGAC GGTGGCCATG  15540
GAAGTCGGAA TCCGCTAAGG AGTGTGTAAC AACTCACCTG CCGAATCAAC TAGCCCTGAA  15600
AATGGATGGC GCTGGAGCGT CGGGCCCATA CCCGGCCGTC GCCGCAGTCG GAACGGAACG  15660
GGACGGGAGC GGCCGCGGGT GCGCGTCTCT CGGGGTCGGG GGTGCGTGGC GGGGGCCCGT  15720
CCCCCGCCTC CCCTCCGCGC GCCGGGTTCG CCCCCGCGGC GTCGGGCCCC GCGGAGCCTA  15780
CGCCGCGACG AGTAGGAGGG CCGCTGCGGT GAGCCTTGAA GCCTAGGGCG CGGGCCCGGG  15840
TGGAGCCGCC GCAGGTGCAG ATCTTGGTGG TAGTAGCAAA TATTCAAACG AGAACTTTGA  15900
AGGCCGAAGT GGAGAAGGGT TCCATGTGAA CAGCAGTTGA ACATGGGTCA GTCGGTCCTG  15960
AGAGATGGGC GAGTGCCGTT CCGAAGGGAC GGGCGATGGC CTCCGTTGCC CTCGGCCGAT  16020
CGAAAGGGAG TCGGGTTCAG ATCCCCGAAT CCGGAGTGGC GGAGATGGGC GCCGCGAGGC  16080
CAGTGCGGTA ACGCGACCGA TCCCGGAGAA GCCGGCGGGA GGCCTCGGGG AGAGTTCTCT  16140
TTTCTTTGTG AAGGGCAGGG CGCCCTGGAA TGGGTTCGCC CCGAGAGAGG GGCCCGTGCC  16200
TTGGAAAGCG TCGCGGTTCC GGCGGCGTCC GGTGAGCTCT CGCTGGCCCT TGAAAATCCG  16260
GGGGAGAGGG TGTAAATCTC GCGCCGGGCC GTACCCATAT CCGCAGCAGG TCTCCAAGGT  16320
GAACAGCCTC TGGCATGTTG GAACAATGTA GGTAAGGGAA GTCGGCAAGC CGGATCCGTA  16380
ACTTCGGGAT AAGGATTGGC TCTAAGGGCT GGGTCGGTCG GGCTGGGGCG CGAAGCGGGG  16440
CTGGGCGCGC GCCGCGGCTG GACGAGGCGC CGCCGCCCTC TCCCACGTCC GGGGAGACCC  16500
CCCGTCCTTT CCGCCCGGGC CCGCCCTCCC CTCTTCCCCG CGGGGCCCCG TCGTCCCCCG  16560
CGTCGTCGCC ACCTCTCTTC CCCCCTCCTT CTTCCCGTCG GGGGGCGGGT CGGGGGTCGG  16620
CGCGCGGCGC GGGCTCCGGG GCGGCGGGTC CAACCCCGCG GGGGTTCCGG AGCGGGAGGA  16680
ACCAGCGGTC CCCGGTGGGG CGGGGGGCCC GGACACTCGG GGGGCCGGCG GCGGCGGCGA  16740
CTCTGGACGC GAGCCGGGCC CTTCCCGTGG ATCGCCTCAG CTGCGGCGGG CGTCGCGGCC  16800
GCTCCCGGGG AGCCCGGCGG GTGCCGGCGC GGGTCCCCTC CCCGCGGGGC CTCGCTCCAC  16860
CCCCCCATCG CCTCTCCCGA GGTGCGTGGC GGGGGCGGGC GGGCGTGTCC CGCGCGTGTG  16920
GGGGGAACCT CCGCGTCGGT GTTCCCCCGC CGGGTCCGCC CCCCGGGCCG CGGTTTTCCG  16980
CGCGGCGCCC CCGCCTCGGC CGGCGCCTAG CAGCCGACTT AGAACTGGTG CGGACCAGGG  17040
GAATCCGACT GTTTAATTAA AACAAAGCAT CGCGAAGGCC CGCGGCGGGT GTTGACGCGA  17100
TGTGATTTCT GCCCAGTGCT CTGAATGTCA AAGTGAAGAA ATTCAATGAA GCGCGGGTAA  17160
ACGGCGGGAG TAACTATGAC TCTCTTAAGG TAGCCAAATG CCTCGTCATC TAATTAGTGA  17220
CGCGCATGAA TGGATGAACG AGATTCCCAC TGTCCCTACC TACTATCCAG CGAAACCACA  17280
GCCAAGGGAA CGGGCTTGGC GGAATCAGCG GGGAAAGAAG ACCCTGTTGA GCTTGACTCT  17340
AGTCTGGCAC GGTGAAGAGA CATGAGAGGT GTAGAATAAG TGGGAGGCCC CCGGCGCCCG  17400
GCCCCGTCCT CGCGTCGGGG TCGGGGCACG CCGGCCTCGC GGGCCGCCGG TGAAATACCA  17460
CTACTCTCAT CGTTTTTTCA CTGACCCGGT GAGGCGGGGG GGCGAGCCCC GAGGGGCTCT  17520
CGCTTCTGGC GCCAAGCGTC CGTCCCGCGC GTGCGGGCGG GCGCGACCCG CTCCGGGGAC  17580
AGTGCCAGGT GGGGAGTTTG ACTGGGGCGG TACACCTGTC AAACGGTAAC GCAGGTGTCC  17640
TAAGGCGAGC TCAGGGAGGA CAGAAACCTC CCGTGGAGCA GAAGGGCAAA AGCTCGCTTG  17700
ATCTTGATTT TCAGTACGAA TACAGACCGT GAAAGCGGGG CCTCACGATC CTTCTGACCT  17760
TTTGGGTTTT AAGCAGGAGG TGTCAGAAAA GTTACCACAG GGATAACTGG CTTGTGGCGG  17820
CCAAGCGTTC ATAGCGACGT CGCTTTTTGA TCCTTCGATG TCGGCTCTTC CTATCATTGT  17880
GAAGCAGAAT TCACCAAGCG TTGGATTGTT CACCCACTAA TAGGGAACGT GAGCTGGGTT  17940
TAGACCGTCG TGAGACAGGT TAGTTTTACC CTACTGATGA TGTGTTGTTG CCATGGTAAT  18000
CCTGCTCAGT ACGAGAGGAA CCGCAGGTTC AGACATTTGG TGTATGTGCT TGGCTGAGGA  18060
GCCAATGGGG CGAAGCTACC ATCTGTGGGA TTATGACTGA ACGCCTCTAA GTCAGAATCC  18120
GCCCAAGCGG AACGATACGG CAGCGCCGAA GGAGCCTCGG TTGGCCCCGG ATAGCCGGGT  18180
CCCCGTCCGT CCCGCTCGGC GGGGTCCCCG CGTCGCCCCG CGGCGGCGCG GGGTCTCCCC  18240
CCGCCGGGCG TCGGGACCGG GGTCCGGTGC GGAGAGCCGT TCGTCTTGGG AAACGGGGTG  18300
CGGCCGGAAA GGGGGCCGCC CTCTCGCCCG TCACGTTGAA CGCACGTTCG TGTGGAACCT  18360
GGCGCTAAAC CATTCGTAGA CGACCTGCTT CTGGGTCGGG GTTTCGTACG TAGCAGAGCA  18420
GCTCCCTCGC TGCGATCTAT TGAAAGTCAG CCCTCGACAC AAGGGTTTGT CTCTGCGGGC  18480
TTTCCCGTCG CACGCCCGCT CGCTCGCACG CGACCGTGTC GCCGCCCGGG CGTCACGGGG  18540
GCGGTCGCCT CGGCCCCCGC GCGGTTGCCC GAACGACCGT GTGGTGGTTG GGGGGGGGAT  18600
CGTCTTCTCC TCCGTCTCCC GAGGACGGTT CGTTTCTCTT TCCCCTTCCG TCGCTCTCCT  18660
TGGGTGTGGG AGCCTCGTGC CGTCGCGACC GCGGCCTGCC GTCGCCTGCC GCCGCAGCCC  18720
CTTGCCCTCC GGCCTTGGCC AAGCCGGAGG GCGGAGGAGG GGGATCGGCG GCGGCGGCGA  18780
```

```
CCGCGGCGCG GTGACGCACG GTGGGATCCC CATCCTCGGC GCGTCCGTCG GGGACGGCCG  18840
GTTGGAGGGG CGGGAGGGGT TTTTCCCGTG AACGCCGCGT TCGGCGCCAG GCCTCTGGCG  18900
GCCGGGGGGG CGCTCTCTCC GCCCGAGCAT CCCCACTCCC GCCCCTCCTC TTCGCGCGCC  18960
GCGGCGGCGA CGTGCGTACG AGGGGAGGAT GTCGCGGTGT GGAGGCGGAG AGGGTCCGGC  19020
GCGGCGCCTC TTCCATTTTT TCCCCCCCAA CTTCGGAGGT CGACCAGTAC TCCGGGCGAC  19080
ACTTTGTTTT TTTTTTTTCC CCCGATGCTG GAGGTCGACC AGATGTCCGA AAGTGTCCCC  19140
CCCCCCCCCC CCCCCCGGCG CGGAGCGGCG GGGCCACTCT GGACTCTTTT TTTTTTTTTT  19200
TTTTTTTTTT TTAAATTCCT GGAACCTTTA GGTCGACCAG TTGTCCGTCT TTTACTCCTT  19260
CATATAGGTC GACCAGTACT CCGGGTGGTA CTTTGTCTTT TTCTGAAAAT CCCAGAGGTC  19320
GACCAGATAT CCGAAAGTCC TCTCTTTCCC TTTACTCTTC CCCACAGCGA TTCTCTTTTT  19380
TTTTTTTTTT TTTGGTGTGC CTCTTTTTGA CTTATATACA TGTAAATAGT GTGTACGTTT  19440
ATATACTTAT AGGAGGAGGT CGACCAGTAC TCCGGGCGAC ACTTTGTTTT TTTTTTTTTT  19500
TCCACCGATG ATGGAGGTCG ACCAGATGTC CGAAAGTGTC CCGTCCCCCC CCTCCCCCCC  19560
CCGCGACGCG GCGGGCTCAC TCTGGACTCT TTTTTTTTTT TTTTTTTTTT TTTAAATTTC  19620
TGGAACCTTA AGGTCGACCA GTTGTCCGTC TTTCACTCAT TCATATAGGT CGACCGGTGG  19680
TACTTTGTCT TTTTCTGAAA ATCGCAGAGG TCGACCAGAT GTCAGAAAGT CTGGTGGTCG  19740
ATAAATTATC TGATCTAGAT TTGTTTTTCT GTTTTTCAGT TTTGTGTTGT TTTGTGTTGT  19800
TTTGTGTTGT TTTGTTTTGT TTTGTTTTGT TTTGTTTTGT TTTGTTTTGT TTTGTTTTGT  19860
TTTGTGTTGT GTTGTGTTGT GTTGTGTTGG GTTGGGTTGG GTTGGGTTGG GTTGGGTTGG  19920
GTTGGGTTGG GTTGGGTTGT GTTGTTTGGT TTTGTGTTGT TTGGTGTTGT TGGTTTTGTT  19980
TTGTTTGCTG TTGTTTTGTG TTTTGCGGGT CGAACAGTTG TCCCTAACCG AGTTTTTTTG  20040
TACACAAACA TGCACTTTTT TTAAAATAAA TTTTTAAAAT AAATGCGAAA ATCGACCAAT  20100
TATCCCTTTC CTTCTCTCTC TTTTTTAAAA ATTTTCTTTG TGTGTGTGTG TGTGTGTGTG  20160
TGTGTGTGTG TGCGTGTGTG TGTGTGTGTG CGTGCAGCGT GCGCGCGCTC GTTTTATAAA  20220
TACTTATAAT AATAGGTCGC CGGGTGGTGG TAGCTTCCCG GACTCCAGAG GCAGAGGCAG  20280
GCAGACTTCT GAGTTCGAGG CCAGCCTGGT CTACAGAGGA ACCCTGTCTC GAAAAATGAA  20340
AATAAATACA TACATACATA CATACATACA TACATACATA CATACATACA TACATATGAG  20400
GTTGACCAGT TGTCAATCCT TTAGAATTTT GTTTTTAATT AATGTGATAG AGAGATAGAT  20460
AATAGATAGA TGGATAGAGT GATACAAATA TAGGTTTTTT TTTCAGTAAA TATGAGGTTG  20520
ATTAACCACT TTTCCCTTTT TAGGTTTTTT TTTTTTTCCC CTGTCCATGT GGTTGCTGGG  20580
ATTTGAACTC AGGACCCTGG CAGGTCAACT GGAAAACGTG TTTTCTATAT ATATAAATAG  20640
TGGTCTGTCT GCTGTTTGTT TGTTTGCTTG CTTGCTTGCT TGCTTGCTTG CTTGCTTGCT  20700
TGCTTTTTTT TTTCTTCTGA GACAGTATTT CTCTGTGTAA CCTGGTGCCC TGAAACTCAC  20760
TCTGTAGACC AGCCTGGCCT CAATCGAACT CAGAAATCCT CCTGCCTCTT GTCTACCTCC  20820
CAATTTTGGA GTAAAGGTGT GCTACACCAC TGCCTGGCAT TATTATCATT ATCATTATTA  20880
ATTTTATTAT TAGACAGAAC GAAATCAACT AGTTGGTCCT GTTTCGTTAA TTCATTTGAA  20940
ATTAGTTGGA CCAATTAGTT GGCTGGTTTG GGAGGTTTCT TTTGTTTCCG ATTTGGGTGT  21000
TTGTGGGGCT GGGGATCAGG TATCTCAACG GAATGCATGA AGGTTAAGGT GAGATGGCTC  21060
GATTTTTGTA AAGATTACTT TTCTTAGTCT GAGGAAAAAA TAAAATAATA TTGGGCTACG  21120
TTTCATTGCT TCATTTCTAT TTCTCTTTCT TTCTTTCTTT CTTTCAGATA AGGAGGTCGG  21180
CCAGTTCCTC CTGCCTTCTG GAAGATGTAG GCATTGCATT GGGAAAAGCA TTGTTTGAGA  21240
GATGTGCTAG TGAACCAGAG AGTTTGGATG TCAAGCCGTA TAATGTTTAT TACAATATAG  21300
AAAAGTTCTA ACAAAGTGAT CTTTAACTTT TTTTTTTTTT TTTCTCCTTC TACTTCTACT  21360
TGTTCTCACT CTGCCACCAA CGCGCTTTGT ACATTGAATG TGAGCTTTGT TTTGCTTAAC  21420
AGACATATAT TTTTTCTTTT GGTTTTGCTT GACATGGTTT CCCTTTCTAT CCGTGCAGGG  21480
TTCCCAGACG GCCTTTTGAG AATAAAATGG GAGGCCAGAA CCAAAGTCTT TTGAATAAAG  21540
CACCACAACT CTAACCTGTT TGGCTGTTTT CCTTCCCAAG GCACAGATCT TTCCCAGCAT  21600
GGAAAAGCAT GTAGCAGTTG TAGGACACAC TAGACGAGAG CACCAGATCT CATTGTGGGT  21660
GGTTGTGAAC CACCCACCAT GTGGTTGCCT GGGATTTGAA CTCAGGATCT TCAGAAGACG  21720
AGTCAGGGCT CTAAACCGAT GAGCCATCTC TCCAGCCCTC CTACATTCCT TCTTAAGGCA  21780
TGAATGATCC CAGCATGGGA AGACAGTCTG CCCTCTTTGT GGTATATCAC CATATACTCA  21840
ATAAAATAAT GAAATGAATG AAGTCTCCAC GTATTTATTT CTTCGAGCTA TCTAAATTCT  21900
CTCACAGCAC CTCCCCCTCC CCCACACTGC CTTTCTCCCT ATGTTTGGGT GGGGCTGGGG  21960
GAGGGGTGGG GTGGGGGCAG GGATCTGCAT GTCTTCTTGC AGGTCTGTGA ACTATTTGCG  22020
ATGGCCTGGT TCTCTGAACT GTTGAGCCTT GTCTATCCAG AGGCTGACTG GCTAGTTTTC  22080
TACCTGAAGT CCCTGAGTGA TGATTTCCCT GTGAATTC                          22118
```

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 42999 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

```
GCTGACACGC TGTCCTCTGG CGACCTGTCG TCGGAGAGGT TGGGCCTCCG GATGCGCGCG     60
GGGCTCTGGC CTCACGGTGA CCGGCTAGCC GGCCGCGCTC CTGCCTTGAG CCGCCTGCCG    120
CGGCCCGCGG GCCTGCTGTT CTCTCGCGCG TCCGAGCGTC CCGACTCCCG GTGCCGGCCC    180
GGGTCCGGGT CTCTGACCCA CCCGGGGGCG GCGGGGAAGG CGGCGAGGGC CACCGTGCCC    240
CGTGCGCTCT CCGCTGCGGG CGCCCGGGGC GCCGCACAAC CCCACCCGCT GGCTCCGTGC    300
CGTGCGTGTC AGGCGTTCTC GTCTCCGCGG GGTTGTCCGC CGCCCCTTCC CCGGAGTGGG    360
GGGTGGCCGG AGCCGATCGG CTCGCTGGCC GGCCGGCCTC CGCTCCCGGG GGGCTCTTCG    420
ATCGATGTGG TGACGTCGTG CTCTCCCGGG CCGGGTCCGA GCCGCGACGG GCGAGGGGCG    480
GACGTTCGTG GCGAACGGGA CCGTCCTTCT CGCTCCGCCC GCGCGGTCCC CTCGTCTGCT    540
CCTCTCCCCG CCCGCCGGCC GGCGTGTGGG AAGGCGTGGG GTGCGGACCC CGGCCCGACC    600
TCGCCGTCCC GCCCGCCGCC TTCGCTTCGC GGGTGCGGGC CGGCGGGGTC CTCTGACGCG    660
GCAGACAGCC CTGCCTGTCG CCTCCAGTGG TTGTCGACTT GCGGGCGGCC CCCCTCCGCG    720
GCGGTGGGGG TGCCGTCCCG CCGGCCCGTC GTGCTGCCCT CTCGGGGGGG GTTTGCGCGA    780
GCGTCGGCTC CGCCTGGGCC CTTGCGGTGC TCCTGGAGCG CTCCGGGTTG TCCCTCAGGT    840
GCCCGAGGCC GAACGGTGGT GTGTCGTTCC CGCCCCCGGC GCCCCCTCCT CCGGTCGCCG    900
CCGCGGTGTC CGCGCGTGGG TCCTGAGGGA GCTCGTCGGT GTGGGGTTCG AGGCGGTTTG    960
AGTGAGACGA GACGAGACGC GCCCCTCCCA CGCGGGGAAG GGCGCCCGCC TGCTCTCGGT   1020
GAGCGCACGT CCCGTGCTCC CCTCTGGCGG GTGCGCGCGG GCCGTGTGAG CGATCGCGGT   1080
GGGTTCGGGC CGGTGTGACG CGTGCGCCGG CCGGCCGCCG AGGGGCTGCC GTTCTGCCTC   1140
CGACCGGTCG TGTGTGGGTT GACTTCGGAG GCGCTCTGCC TCGGAAGGAA GGAGGTGGGT   1200
GGACGGGGGG GCCTGGTGGG GTTGCGCGCA CGCGCGCACC GGCCGGGCCC CCGCCCTGAA   1260
CGCGAACGCT CGAGGTGGCC GCGCGCAGGT GTTTCCTCGT ACCGCAGGGC CCCCTCCCTT   1320
CCCCAGGCGT CCCTCGGCGC CTCTGCGGGC CCGAGGAGGA GCGGCTGGCG GGTGGGGGGA   1380
GTGTGACCCA CCCTCGGTGA GAAAAGCCTT CTCTAGCGAT CTGAGAGGCG TGCCTTGGGG   1440
GTACCGGATC CCCCGGGCCG CCGCCTCTGT CTCTGCCTCC GTTATGGTAG CGCTGCCGTA   1500
GCGACCCGCT CGCAGAGGAC CCTCCTCCGC TTCCCCCTCG ACGGGGTTGG GGGGGAGAAG   1560
CGAGGGTTCC GCCGGCCACC GCGGTGGTGG CCGAGTGCGG CTCGTCGCCT ACTGTGGCCC   1620
GCGCCTCCCC CTTCCGAGTC GGGGGAGGAT CCCGCCGGGC CGGGCCCGGC GCTCCCACCC   1680
AGCGGGTTGG GACGCGGCGG CCGGCGGGCG GTGGGTGTGC GCGCCCGGCG CTCTGTCCGG   1740
CGCGTGACCC CCTCCGTCCG CGAGTCGGCT CTCCGCCCGC TCCCGTGCCG AGTCGTGACC   1800
GGTGCCGACG ACCGCGTTTG CGTGGCACGG GGTCGGGCCC GCCTGGCCCT GGGAAAGCGT   1860
CCCACGGTGG GGGCGCGCCG GTCTCCCGGA GCGGGACCGG GTCGGAGGAT GGACGAGAAT   1920
CACGAGCGAC GGTGGTGGTG GCGTGTCGGG TTCGTGGCTG CGGTCGCTCC GGGGCCCCCG   1980
GTGGCGGGGC CCCGGGGCTC GCGAGGCGGT TCTCGGTGGG GGCCGAGGGC CGTCCGGCGT   2040
CCCAGGCGGG GCGCCGCGGG ACCGCCCTCG TGTCTGTGGC GGTGGGATCC CGCGGCCGTG   2100
TTTTCCTGGT GGCCCGGCCG TGCCTGAGGT TTCTCCCCGA GCCGCCGCCT CTGCGGGCTC   2160
CCGGGTGCCC TTGCCCTCGC GGTCCCCGGC CCTCGCCCGT CTGTGCCCTC TTCCCCGCCC   2220
GCCGCCCGCC GATCCTCTTC TTCCCCCCGA GCGGCTCACC GGCTTCACGT CCGTTGGTGG   2280
CCCCGCCTGG GACCGAACCC GGCACCGCCT CGTGGGGCGC CGCCGCCGGC CACTGATCGG   2340
CCCGGCGTCC GCGTCCCCCG GCGCGCGCCT TGGGGACCGG GTCGGTGGCG CGCCGCGTGG   2400
GGCCCGGTGG GCTTCCCGGA GGGTTCCGGG GGTCGGCCTG CGGCGCGTGC GGGGGAGGAG   2460
ACGGTTCCGG GGGACCGGCC GCGGCTGCGG CGGCGGCGGT GGTGGGGGGA GCCGCGGGGA   2520
TCGCCGAGGG CCGGTCGGCC GCCCCGGGTG CCCCGCGGTG CCGCCGGCGG CGGTGAGGCC   2580
CCGCGCGTGT GTCCCGGCTG CGGTCGGCCG CGCTCGAGGG GTCCCCGTGG CGTCCCCTTC   2640
CCCGCCGGCC GCCTTTCTCG CGCCTTCCCC GTCGCCCCGG CCTCGCCCGT GGTCTCTCGT   2700
CTTCTCCCGG CCCGCTCTTC CGAACCGGGT CGGCGCGTCC CCCGGGTGCG CCTCGCTTCC   2760
CGGGCCTGCC GCGGCCCTTC CCCGAGGCGT CCGTCCCGGG CGTCGGCGTC GGGGAGAGCC   2820
CGTCCTCCCC GCGTGGCGTC GCCCCGTTCG GCGCGCGCGT GCGCCCGAGC GCGGCCCGGT   2880
GGTCCCTCCC GGACAGGCGT TCGTGCGACG TGTGGCGTGG GTCGACCTCC GCCTTGCCGG   2940
TCGCTCGCCC TCTCCCCGGG TCGGGGGGTG GGGCCCGGGC CGGGGCCTCG GCCCCGGTCG   3000
CTGCCTCCCG TCCCGGGCGG GGGCGGGCGC GCCGGCCGGC CTCGGTCGCC CTCCCTTGGC   3060
CGTCGTGTGG CGTGTGCCAC CCCTGCGCCG GCGCCCGCCG GCGGGGCTCG GAGCCGGGCT   3120
TCGGCCGGGC CCCGGGCCCT CGACCGGACC GGCTGCGCGG GCGCTGCGGC CGCACGGCGC   3180
GACTGTCCCC GGGCCGGGCA CCGCGGTCCG CCTCTCGCTC GCCGCCCGGA CGTCGGGGCC   3240
GCCCCGCGGG GCGGGCGGAG CGCCGTCCCC GCCTCGCCGC CGCCCGCGGG CGCCGGCCGC   3300
GCGCGCGCGC GCGTGGCCGC CGGTCCCTCC CGGCCGCCGG GCGCGGGTCG GGCCGTCCGC   3360
CTCCTCGCGG GCGGGCGCGA CGAAGAAGCG TCGCGGGTCT GTGGCGCGGG GCCCCCGGTG   3420
```

```
GTCGTGTCGC GTGGGGGGCG GGTGGTTGGG GCGTCCGGTT CGCCGCGCCC CGCCCCGGCC   3480
CCACCGGTCC CGGCCGCCGC CCCCGCGCCC GCTCGCTCCC TCCCGTCCGC CCGTCCGCGG   3540
CCCGTCCGTC CGTCCGTCCG TCGTCCTCCT CGCTTGCGGG GCGCCGGGCC CGTCCTCGCG   3600
AGGCCCCCCG GCCGGCCGTC CGGCCGCGTC GGGGGCTCGC CGCGCTCTAC CTTACCTACC   3660
TGGTTGATCC TGCCAGTAGC ATATGCTTGT CTCAAAGATT AAGCCATGCA TGTCTAAGTA   3720
CGCACGGCCG GTACAGTGAA ACTGCGAATG GCTCATTAAA TCAGTTATGG TTCCTTTGGT   3780
CGCTCGCTCC TCTCCTACTT GGATAACTGT GGTAATTCTA GAGCTAATAC ATGCCGACGG   3840
GCGCTGACCC CCTTCGCGGG GGGGATGCGT GCATTTATCA GATCAAAACC AACCCGGTCA   3900
GCCCCTCTCC GGCCCCGGCC GGGGGGCGGG CGCCGGCGGC TTTGGTGACT CTAGATAACC   3960
TCGGGCCGAT CGCACGCCCC CCGTGGCGGC GACGACCCAT TCGAACGTCT GCCCTATCAA   4020
CTTTCGATGG TAGTCGCCGT GCCTACCATG GTGACCACGG GTGACGGGGA ATCAGGGTTC   4080
GATTCCGGAG AGGGAGCCTG AGAAACGGCT ACCACATCCA AGGAAGGCAG CAGGCGCGCA   4140
AATTACCCAC TCCCGACCCG GGGAGGTAGT GACGAAAAAT AACAATACAG GACTCTTTCG   4200
AGGCCCTGTA ATTGGAATGA GTCCACTTTA AATCCTTTAA CGAGGATCCA TTGGAGGGCA   4260
AGTCTGGTGC CAGCAGCCGC GGTAATTCCA GCTCCAATAG CGTATATTAA AGTTGCTGCA   4320
GTTAAAAAGC TCGTAGTTGG ATCTTGGGAG CGGGCGGGCG GTCCGCCGCG AGGCGAGCCA   4380
CCGCCCGTCC CCGCCCCTTG CCTCTCGGCG CCCCCTCGAT GCTCTTAGCT GAGTGTCCCG   4440
CGGGGCCCGA AGCGTTTACT TTGAAAAAAT TAGAGTGTTC AAAGCAGGCC CGAGCCGCCT   4500
GGATACCGCA GCTAGGAATA ATGGAATAGG ACCGCGGTTC TATTTTGTTG GTTTTCGGAA   4560
CTGAGGCCAT GATTAAGAGG GACGGCCGGG GGCATTCGTA TTGCGCCGCT AGAGGTGAAA   4620
TTCTTGGACC GGCGCAAGAC GGACCAGAGC GAAAGCATTT GCCAAGAATG TTTTCATTAA   4680
TCAAGAACGA AAGTCGGAGG TTCGAAGACG ATCAGATACC GTCGTAGTTC CGACCATAAA   4740
CGATGCCGAC CGGCGATGCG GCGGCGTTAT TCCCATGACC CGCCGGGCAG CTTCCGGGAA   4800
ACCAAAGTCT TTGGGTTCCG GGGGGAGTAT GGTTGCAAAG CTGAAACTTA AAGGAATTGA   4860
CGGAAGGGCA CCACCAGGAG TGGAGCCTGC GGCTTAATTT GACTCAACAC GGGAAACCTC   4920
ACCCGGCCCG GACACGGACA GGATTGACAG ATTGATAGCT CTTTCTCGAT TCCGTGGGTG   4980
GTGGTGCATG GCCGTTCTTA GTTGGTGGAG CGATTTGTCT GGTTAATTCC GATAACGAAC   5040
GAGACTCTGG CATGCTAACT AGTTACGCGA CCCCCGAGCG GTCGGCGTCC CCCAACTTCT   5100
TAGAGGGACA AGTGGCGTTC AGCCACCCGA GATTGAGCAA TAACAGGTCT GTGATGCCCT   5160
TAGATGTCCG GGGCTGCACG CGCGCTACAC TGACTGGCTC AGCGTGTGCC TACCCTACGC   5220
CGGCAGGCGC GGGTAACCCG TTGAACCCCA TTCGTGATGG GGATCGGGGA TTGCAATTAT   5280
TCCCCATGAA CGAGGGAATT CCCGAGTAAG TGCGGGTCAT AAGCTTGCGT TGATTAAGTC   5340
CCTGCCCTTT GTACACACCG CCCGTCGCTA CTACCGATTG GATGGTTTAG TGAGGCCCTC   5400
GGATCGGCCC CGCCGGGGTC GGCCCACGGC CCTGGCGGAG CGCTGAGAAG ACGGTCGAAC   5460
TTGACTATCT AGAGGAAGTA AAAGTCGTAA CAAGGTTTCC GTAGGTGAAC CTGCGGAAGG   5520
ATCATTAACG GAGCCCGGAG GGCGAGGCCC GCGGCGGCGC CGCCGCCGCC GCGCGCTTCC   5580
CTCCGCACAC CCACCCCCCC ACCGCGACGC GGCGCGTGCG CGGGCGGGGC CCGCGTGCCC   5640
GTTCGTTCGC TCGCTCGTTC GTTCGCCGCC CGGCCCCGCC GCCGCGAGAG CCGAGAACTC   5700
GGGAGGGAGA CGGGGGGGAG AGAGAGAGAG AGAGAGAGAG AGAGAGAGAG AGAGAGAGAA   5760
AGAAGGGCGT GTCGTTGGTG TGCGCGTGTC GTGGGGCCGG CGGGCGGCGG GGAGCGGTCC   5820
CCGGCCGCGG CCCCGACGAC GTGGGTGTCG GCGGGCGCGG GGGCGGTTCT CGGCGGCGTC   5880
GCGGCGGGTC TGGGGGGGTC TCGGTGCCCT CCTCCCCGCC GGGGCCCGTC GTCCGGCCCC   5940
GCCGCGCCGG CTCCCCGTCT TCGGGGCCGG CCGGATTCCC GTCGCCTCCG CCGCGCCGCT   6000
CCGCGCCGCC GGGCACGGCC CCGCTCGCTC TCCCCGGCCT TCCCGCTAGG GCGTCTCGAG   6060
GGTCGGGGGC CGGACGCCGG TCCCCTCCCC CGCCTCCTCG TCCGCCCCCC CGCCGTCCAG   6120
GTACCTAGCG CGTTCCGGCG CGGAGGTTTA AAGACCCCTT GGGGGGATCG CCCGTCCGCC   6180
CGTGGGTCGG GGGCGGTGGT GGGCCCGCGG GGGAGTCCCG TCGGGAGGGG CCCGGCCCCT   6240
CCCGCGCCTC CACCGCGGAC TCCGCTCCCC GGCCGGGGCC GCGCCGCCGC CGCCGCCGCG   6300
GCGGCCGTCG GGTGGGGGCT TTACCCGGCG GCCGTCGCGC GCCTGCCGCG CGTGTGGCGT   6360
GCGCCCCGCG CCGTGGGGGC GGGAACCCCC GGGCGCCTGT GGGGTGGTGT CCGCGCTCGC   6420
CCCCGCGTGG GCGGCGCGCG CCTCCCCGTG GTGTGAAACC TTCCGACCCC TCTCCGGAGT   6480
CCGGTCCCGT TTGCTGTCTC GTCTGGCCGG CCTGAGGCAA CCCCCTCTCC TCTTGGGCGG   6540
GGGGGGCGGG GGACGTGCC  GCGCCAGGAA GGGCCTCCTC CCGGTGCGTC GTCGGGAGCG   6600
CCCTCGCCAA ATCGACCTCG TACGACTCTT AGCGGTGGAT CACTCGGCTC GTGCGTCGAT   6660
GAAGAACGCA GCTAGCTGCG AGAATTAATG TGAATTGCAG GACACATTGA TCATCGACAC   6720
TTCGAACGCA CTTGCGGCCC CGGGTTCCTC CCGGGGCTAC GCCTGTCTGA GCGTCGCTTG   6780
CCGATCAATC GCCCCGGGGG TGCCTCCGGG CTCCTCGGGG TGCGCGGCTG GGGGTTCCCT   6840
CGCAGGGCCC GCCGGGGGCC CTCCGTCCCC CTAAGCGCAG ACCCGGCGGC GTCCGCCCTC   6900
CTCTTGCCGC CGCGCCCGCC CCTTCCCCCT CCCCCCGCGG GCCCTGCGTG GTCACGCGTC   6960
GGGTGGCGGG GGGGAGAGGG GGGCGCGCCC GGCTGAGAGA GACGGGGAGG GCGGCGCCGC   7020
CGCCGGAAGA CGGAGAGGGA AAGAGAGAGC CGGCTCGGGC CGAGTTCCCG TGGCCGCCGC   7080
CTGCGGTCCG GGTTCCTCCC TCGGGGGGCT CCCTCGCGCC GCGCGCGGCT CGGGGTTCGG   7140
GGTTCGTCGG CCCCGGCCGG GTGGAAGGTC CCGTGCCCGT CGTCGTCGTC GTCGCGCGTC   7200
GTCGGCGGTG GGGGCGTGTT GCGTGCGGTG TGGTGGTGGG GGAGGAGGAA GGCGGGTCCG   7260
GAAGGGGAAG GGTGCCGGCG GGGAGAGAGG GTCGGGGGAG CGCGTCCCGG TCGCCGCGGT   7320
```

```
TCCGCCGCCC GCCCCCGGTG GCGGCCCGGC GTCCGGCCGA CCGGCCGCTC CCCGCGCCCC   7380
TCCTCCTCCC CGCCGCCCCT CCTCCGAGGC CCCGCCCGTC CTCCTCGCCC TCCCCGCGCG   7440
TACGCGCGCG CGCCCGCCCG CCCGGCTCGC CTCGCGGCGC GTCGGCCGGG GCCGGGAGCC   7500
CGCCCCGCCG CCCGCCCGTG GCCGCGGCGC CGGGGTTCGC GTGTCCCCGG CGGCGACCCG   7560
CGGGACGCCG CGGTGTCGTC CGCCGTCGCG CGCCCGCCTC CGGCTCGCGG CCGCGCCGCG   7620
CCGCGCCGGG GCCCCGTCCC GAGCTTCCGC GTCGGGGCGG CGCGGCTCCG CCGCCGCGTC   7680
CTCGGACCCG TCCCCCCGAC CTCCGCGGGG GAGACGCGCC GGGGCGTGCG GCGCCCGTCC   7740
CGCCCCCGGC CCGTGCCCCT CCCTCCGGTC GTCCCGCTCC GGCGGGGCGG CGCGGGGGCG   7800
CCGTCGGCCG CGCGCTCTCT CTCCCGTCGC CTCTCCCCCT CGCCGGGCCC GTCTCCCGAC   7860
GGAGCGTCGG GCGGGCGGTC GGGCCGGCGC GATTCCGTCC GTCCGTCCGC CGAGCGGCCC   7920
GTCCCCCTCC GAGACGCGAC CTCAGATCAG ACGTGGCGAC CCGCTGAATT TAAGCATATT   7980
AGTCAGCGGA GGAAAAGAAA CTAACCAGGA TTCCCTCAGT AACGGCGAGT GAACAGGGAA   8040
GAGCCCAGCG CCGAATCCCC GCCCCGCGGG GCGCGGGACA TGTGGCGTAC GGAAGACCCG   8100
CTCCCCGGCG CCGCTCGTGG GGGGCCCAAG TCCTTCTGAT CGAGGCCCAG CCCGTGGACG   8160
GTGTGAGGCC GGTAGCGGCC GGCGCGCGCC CGGGTCTTCC CGGAGTCGGG TTGCTTGGGA   8220
ATGCAGCCCA AAGCGGGTGG TAAACTCCAT CTAAGGCTAA ATACCGGCAC GAGACCGATA   8280
GTCAACAAGT ACCGTAAGGG AAAGTTGAAA AGAACTTTGA AGAGAGAGTT CAAGAGGGCG   8340
TGAAACCGTT AAGAGGTAAA CGGGTGGGGT CCGCGCAGTC CGCCCGGAGG ATTCAACCCG   8400
GCGGCGGGTC CGGCCGTGTC GGCGGCCCGG CGGATCTTTC CCGCCCCCCG TTCCTCCCGA   8460
CCCCTCCACC CGCCCTCCCT TCCCCCGCCG CCCCTCCTCC TCCTCCCCGG AGGGGGCGGG   8520
CTCCGGCGGG TGCGGGGGTG GGCGGGCGGG GCCGGGGGTG GGGTCGGCGG GGGACCGTCC   8580
CCCGACCGGC GACCGGCCGC CGCCGGGCGC ATTTCCACCG CGGCGGTGCG CCGCGACCGG   8640
CTCCGGGACG GCTGGGAAGG CCCGGCGGGG AAGGTGGCTC GGGGGGCCCC GTCCGTCCGT   8700
CCGTCCTCCT CCTCCCCCGT CTCCGCCCCC CGGCCCCGCG TCCTCCCTCG GGAGGGCGCG   8760
CGGGTCGGGG CGGCGGCGGC GGCGGCGGTG GCGGCGGCGG CGGGGGCGGC GGGACCGAAA   8820
CCCCCCCCGA GTGTTACAGC CCCCCCGGCA GCAGCACTCG CCGAATCCCG GGGCCGAGGG   8880
AGCGAGACCC GTCGCCGCGC TCTCCCCCCT CCCGGCGCCC ACCCCCGCGG GGAATCCCCC   8940
GCGAGGGGGG TCTCCCCCGC GGGGGCGCGC CGGCGTCTCC TCGTGGGGGG GCCGGGCCAC   9000
CCCTCCCACG GCGCGACCGC TCTCCCACCC CTCCTCCCCG CGCCCCCGCC CCGGCGACGG   9060
GGGGGGTGCC GCGCGCGGGT CGGGGGGCGG GGCGGACTGT CCCCAGTGCG CCCCGGGCGG   9120
GTCGCGCCGT CGGGCCCGGG GGAGGTTCTC TCGGGGCCAC GCGCGCGTCC CCCGAAGAGG   9180
GGGACGGCGG AGCGAGCGCA CGGGGTCGGC GGCGACGTCG GCTACCCACC CGACCCGTCT   9240
TGAAACACGG ACCAAGGAGT CTAACACGTG CGCGAGTCGG GGGCTCGCAC GAAAGCCGCC   9300
GTGGCGCAAT GAAGGTGAAG GCCGGCGCGC TCGCCGGCCG AGGTGGGATC CCGAGGCCTC   9360
TCCAGTCCGC CGAGGGCGCA CCACCGGCCC GTCTCGCCCG CCGCGCCGGG GAGGTGGAGC   9420
ACGAGCGCAC GTGTTAGGAC CCGAAAGATG GTGAACTATG CCTGGGCAGG GCGAAGCCAG   9480
AGGAAACTCT GGTGGAGGTC CGTAGCGGTC CTGACGTGCA AATCGGTCGT CCGACCTGGG   9540
TATAGGGGCG AAAGACTAAT CGAACCATCT AGTAGCTGGT TCCCTCCGAA GTTTCCCTCA   9600
GGATAGCTGG CGCTCTCGCA GACCCGACGC ACCCCCGCCA CGCAGTTTTA TCCGGTAAAG   9660
CGAATGATTA GAGGTCTTGG GGCCGAAACG ATCTCAACCT ATTCTCAAAC TTTAAATGGG   9720
TAAGAAGCCC GGCTCGCTGG CGTGGAGCCG GGCGTGGAAT GCGAGTGCCT AGTGGGCCAC   9780
TTTTGGTAAG CAGAACTGGC GCTGCGGGAT GAACCGAACG CCGGGTTAAG GCGCCCGATG   9840
CCGACGCTCA TCAGACCCCA GAAAAGGTGT TGGTTGATAT AGACAGCAGG ACGGTGGCCA   9900
TGGAAGTCGG AATCCGCTAA GGAGTGTGTA ACAACTCACC TGCCGAATCA ACTAGCCCTG   9960
AAAATGGATG GCGCTGGAGC GTCGGGCCCA TACCCGGCCG TCGCCGGCAG TCGAGAGTGG  10020
ACGGGAGCGG CGGGGGCGGC GCGCGCGCGC GCGCGTGTGG TGTGCGTCGG AGGGCGGCGG  10080
CGGCGGCGGC GGCGGGGGTG TGGGGTCCTT CCCCCGCCCC CCCCCCCACG CCTCCTCCCC  10140
TCCTCCCGCC CACGCCCCGC TCCCCGCCCC CGGAGCCCCG CGGACGCTAC GCCGCGACGA  10200
GTAGGAGGGC CGCTGCGGTG AGCCTTGAAG CCTAGGGCGC GGGCCCGGGT GGAGCCGCCG  10260
CAGGTGCAGA TCTTGGTGGT AGTAGCAAAT ATTCAAACGA GAACTTTGAA GGCCGAAGTG  10320
GAGAAGGGTT CCATGTGAAC AGCAGTTGAA CATGGGTCAG TCGGTCCTGA GAGATGGGCG  10380
AGCGCCGTTC CGAAGGGACG GGCGATGGCC TCCGTTGCCC TCGGCCGATC GAAAGGGAGT  10440
CGGGTTCAGA TCCCCGAATC CGGAGTGGCG GAGATGGGCG CCGCGAGGCG TCCAGTGCGG  10500
TAACGCGACC GATCCCGGAG AAGCCGGCGG GAGCCCCGGG GAGAGTTCTC TTTTCTTTGT  10560
GAAGGGCAGG GCGCCCTGGA ATGGGTTCGC CCCGAGAGAG GGGCCCGTGC CTTGGAAAGC  10620
GTCGCGGTTC CGGCGGCGTC CGGTGAGCTC TCGCTGGCCC TTGAAAATCC GGGGGAGAGG  10680
GTGTAAATCT CGCGCCGGGC CGTACCCATA TCCGCAGCAG GTCTCCAAGG TGAACAGCCT  10740
CTGGCATGTT GGAACAATGT AGGTAAGGGA AGTCGGCAAG CCGGATCCGT AACTTCGGGA  10800
TAAGGATTGG CTCTAAGGGC TGGGTCGGTC GGGCTGGGGC GCGAAGCGGG GCTGGGCGCG  10860
CGCCGCGGCT GGACGAGGCG CGCGCCCCCC CCACGCCCGG GGCACCCCCC TCGCGGCCCT  10920
CCCCCGCCCC ACCCGCGCGC GCCGCTCGCT CCCTCCCCAC CCCGCGCCCT CTCTCTCTCT  10980
CTCTCCCCCG CTCCCCGTCC TCCCCCCTCC CCGGGGGAGC GCCGCGTGGG GGCGCGGCGG  11040
GGGGAGAAGG GTCGGGGCGG CAGGGGCCGC GCGGCGGCCG CCGGGGCGGC CGGCGGGGGC  11100
AGGTCCCCGC GAGGGGGGCC CCGGGGACCC GGGGGGCCGG CGGCGGCGCG GACTCTGGAC  11160
GCGAGCCGGG CCCTTCCCGT GGATCGCCCC AGCTGCGGCG GGCGTCGCGG CCGCCCCCGG  11220
```

```
GGAGCCCGGC GGCGGCGCGG CGCGCCCCCC ACCCCCACCC CACGTCTCGG TCGCGCGCGC  11280
GTCCGCTGGG GGCGGGAGCG GTCGGGCGGC GGCGGTCGGC GGGCGGCGGG GCGGGGCGGT  11340
TCGTCCCCCC GCCCTACCCC CCCGGCCCCG TCCGCCCCCC GTTCCCCCCT CCTCCTCGGC  11400
GCGCGGCGGC GGCGGCGGCA GGCGGCGGAG GGGCCGCGGG CCGGTCCCCC CCGCCGGGTC  11460
CGCCCCCGGG GCCGCGGTTC CGCGCGCGCC TCGCCTCGGC CGGCGCCTAG CAGCCGACTT  11520
AGAACTGGTG CGGACCAGGG GAATCCGACT GTTTAATTAA AACAAAGCAT CGCGAAGGCC  11580
CGCGGCGGGT GTTGACGCGA TGTGATTTCT GCCCAGTGCT CTGAATGTCA AAGTGAAGAA  11640
ATTCAATGAA GCGCGGGTAA ACGGCGGGAG TAACTATGAC TCTCTTAAGG TAGCCAAATG  11700
CCTCGTCATC TAATTAGTGA CGCGCATGAA TGGATGAACG AGATTCCCAC TGTCCCTACC  11760
TACTATCCAG CGAAACCACA GCCAAGGGAA CGGGCTTGGC GGAATCAGCG GGGAAAGAAG  11820
ACCCTGTTGA GCTTGACTCT AGTCTGGCAC GGTGAAGAGA CATGAGAGGT GTAGAATAAG  11880
TGGGAGGCCC CCGGCGCCCC CCCGGTGTCC CCGCGAGGGG CCCGGGGCGG GGTCCGCGGC  11940
CCTGCGGGCC GCCGGTGAAA TACCACTACT CTGATCGTTT TTTCACTGAC CCGGTGAGGC  12000
GGGGGGGCGA GCCCGAGGGG CTCTCGCTTC TGGCGCCAAG CGCCCGCCCG GCCGGGCGCG  12060
ACCCGCTCCG GGGACAGTGC CAGGTGGGGA GTTTGACTGG GGCGGTACAC CTGTCAAACG  12120
GTAACGCAGG TGTCCTAAGG CGAGCTCAGG GAGGACAGAA ACCTCCCGTG GAGCAGAAGG  12180
GCAAAAGCTC GCTTGATCTT GATTTTCAGT ACGAATACAG ACCGTGAAAG CGGGGCCTCA  12240
CGATCCTTCT GACCTTTTGG GTTTTAAGCA GGAGGTGTCA GAAAAGTTAC CACAGGGATA  12300
ACTGGCTTGT GGCGGCCAAG CGTTCATAGC GACGTCGCTT TTTGATCCTT CGATGTCGGC  12360
TCTTCCTATC ATTGTGAAGC AGAATTCGCC AAGCGTTGGA TTGTTCACCC ACTAATAGGG  12420
AACGTGAGCT GGGTTTAGAC CGTCGTGAGA CAGGTTAGTT TTACCCTACT GATGATGTGT  12480
TGTTGCCATG GTAATCCTGC TCAGTACGAG AGGAACCGCA GGTTCAGACA TTTGGTGTAT  12540
GTGCTTGGCT GAGGAGCCAA TGGGGCGAAG CTACCATCTG TGGGATTATG ACTGAACGCC  12600
TCTAAGTCAG AATCCGCCC AGGCGAACGA TACGGCAGCG CCGCGGAGCC TCGGTTGGCC  12660
TCGGATAGCC GGTCCCCCGC CTGTCCCCGC CGGCGGGCCG CCCCCCCCTC CACGCGCCCC  12720
GCCGCGGGAG GGCGCGTGCC CCGCCGCGCG CCGGGACCGG GGTCCGGTGC GGAGTGCCCT  12780
TCGTCCTGGG AAACGGGGCG CGGCCGGAAA GGCGGCCGCC CCCTCGCCCG TCACGCACCG  12840
CACGTTCGTG GGGAACCTGG CGCTAAACCA TTCGTAGACG ACCTGCTTCT GGGTCGGGGT  12900
TTCGTACGTA GCAGAGCAGC TCCCTCGCTG CGATCTATTG AAAGTCAGCC CTCGACACAA  12960
GGGTTTGTCC GCGCGCGCGT GCGTGCGGGG GGCCCGGCGG GCGTGCGCGT TCGGCGCCGT  13020
CCGTCCTTCC GTTCGTCTTC CTCCCTCCCG GCCTCTCCCG CCGACCGCGG CGTGGTGGTG  13080
GGGTGGGGGG GAGGGCGCGC GACCCCGGTC GGCCGCCCCG CTTCTTCGGT TCCCGCCTCC  13140
TCCCCGTTCA CGCCGGGGCG GCTCGTCCGC TCCGGGCCGG GACGGGGTCC GGGGAGCGTG  13200
GTTTGGGAGC CGCGGAGGCG CCGCGCCGAG CCGGGCCCCG TGGCCCGCCG GTCCCCGTCC  13260
CGGGGGTTGG CCGCGCGGCG CGGTGGGGGG CCACCCGGGG TCCCGGCCCT CGCGCGTCCT  13320
TCCTCCTCGC TCCTCCGCAC GGGTCGACCG ACGAACCGCG GGTGGCGGGC GGCGGGCGGC  13380
GAGCCCCACG GGCGTCCCCG CACCCGGCCG ACCTCCGCTC GCGACCTCTC CTCGGTCGGG  13440
CCTCCGGGGT CGACCGCCTG CGCCCGCGGG CGTGAGACTC AGCGGCGTCT CGCCGTGTCC  13500
CGGGTCGACC GCGGCCTTCT CCACCGAGCG GCGGTGTAGG AGTGCCCGTC GGGACGAACC  13560
GCAACCGGAG CGTCCCCGTC TCGGTCGGCA CCTCCGGGGT CGACCAGCTG CCGCCCGCGA  13620
GCTCCGGACT TAGCCGGCGT CTGCACGTGT CCCGGGTCGA CCAGCAGGCG GCCGCCGGAC  13680
GCAGCGGCGC ACGCACGCGA GGGCGTCGAT TCCCCTTCGC GCGCCCGCGC CTCCACCGGC  13740
CTCGGCCCGC GGTGGAGCTG GGACCACGCG GAACTCCCTC TCCCACATTT TTTTCAGCCC  13800
CACCGCGAGT TTGCGTCCGC GGGACCTTTA AGAGGGAGTC ACTGCTGCCG TCAGCCAGTA  13860
CTGCCTCCTC CTTTTTCGCT TTTAGGTTTT GCTTGCCTTT TTTTTTTTTT TTTTTTTTTT  13920
TTTTTTCTTT CTTTCTTTCT TTCTTTCTTT CTTTCTTTCT TTCTTTCTTT CGCTTGTCTT  13980
CTTCTTGTGT TCTCTTCTTG CTCTTCCTCT GTCTGTCTCT CTCTCTCTCT CTCTCTCTGT  14040
CTCTCGCTCT CGCCCTCTCT CTCTTCTCTC TCTCTCTCTC TCTCTCTCTG TCTCTCGCTC  14100
TCGCCCTCTC TCTCTCTCTT CTCTCTGTCT CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT  14160
GTCGCTCTCG CCCTCTCGCT CTCTCTCTGT CTCTGTCTGT GTCTCTCTCT CTCCCTCCCT  14220
CCCTCCCTCC CTCCCTCCCT CCCTCCCCTT CCTTGGCGCC TTCTCGGCTC TTGAGACTTA  14280
GCCGCTGTCT CGCCGTACCC CGGGTCGACC GGCGGGCCTT CTCCACCGAG CGGCGTGCCA  14340
CAGTGCCCGT CGGGACGAGC CGGACCCGCC GCGTCCCCGT CTCGGTCGGC ACCTCCGGGG  14400
TCGACCAGCT GCCGCCCGCG AGCTCCGGAC TTAGCCGGCG TCTGCACGTG TCCCGGGTCG  14460
ACCAGCAGGC GGCCGCCGGA CGCAGCGGCG CACCGACGGA GGGCGCTGAT TCCCGTTCAC  14520
GCGCCCGCGC CTCCACCGGC CTCGGCCCGC CGTGGAGCTG GGACCACGCG GAACTCCCTC  14580
TCCTACATTT TTTTCAGCCC CACCGCGAGT TTGCGTCCGC GGGACCTTTA AGAGGGAGTC  14640
ACTGCTGCCG TCAGCCAGTA CTGCCTCCTC CTTTTTCGCT TTTAGGTTTT GCTTGCCTTT  14700
TTTTTTTTTT TTTTTTTTTT TTTTTTCTTT CTTTCTTTCT TTCTTTCTTT CTTTCTTTCT  14760
TTCTTTCTTT CTTTCGCTCT CGCTCTCTCG CTCTCTCCCT CGCTCGTTTC TTTCTTTCTC  14820
TTTCTCTCTC TCTCTCTCTC TCTCTCTCTC TCTGTCTCTC GCTCTCGCCC TCTCTCTCTC  14880
TTTCTCTCTC TCTCTGTCTC TCTCTCTCTC TCTCTCTCTC TCTCTCTCTC CCTCCCTCCC  14940
TCCCCCTCCC TCCCTCTCTC CCCTTCCTTG GCGCCTTCTC GGCTCTTGAG ACTTAGCCGC  15000
TGTCTCGCCG TGTCCCGGGT CGACCGGCGG GCCTTCTCCA CCGAGCGGCG TGCCACAGTG  15060
CCCGTCGGGA CGAGCCGGAC CCGCCGCGTC CCCGTCTCGG TCGGCACCTC CGGGGTCGAC  15120
```

```
CAGCTGCCGC CCGCGAGCTC CGGACTTAGC CGGCGTCTGC ACGTGTCCCG GGTCGACCAG  15180
CAGGCGGCCG CCGGACGCTG CGGCGCACCG ACGCGAGGGC GTCGATTCCG GTTCACGCGC  15240
CGGCGACCTC CACCGGCCTC GGCCCGCGGT GGAGCTGGGA CCACGCGGAA CTCCCTCTCC  15300
CACATTTTTT TCAGCCCCAC CGCGAGTTTG CGTCCGCGGG ACTTTTAAGA GGGAGTCACT  15360
GCTGCCGTCA GCCAGTAATG CTTCCTCCTT TTTTGCTTTT TGGTTTTGCC TTGCGTTTTC  15420
TTTCTTTCTT TCTTTCTTTC TTTCTTTCTT TCTTTCTTTC TCTCTCTCTC TCTCTCTCTC  15480
TCTCTGTCTC TCTCTCTCTG TCTCTCTCCC CTCCCTCCCT CCTTGGTGCC TTCTCGGCTC  15540
GCTGCTGCTG CTGCCTCTGC CTCCACGGTT CAAGCAAACA GCAAGTTTTC TATTTCGAGT  15600
AAAGACGTAA TTTCACCATT TTGGCCGGGC TGGTCTCGAA CTCCCGACCT AGTGATCCGC  15660
CCGCCTCGGC CTCCCAAAGA CTGCTGGGAG TACAGATGTG AGCCACCATG CCCGGCCGAT  15720
TCCTTCCTTT TTTCAATCTT ATTTTCTGAA CGCTGCCGTG TATGAACATA CATCTACACA  15780
CACACACACA CACACACACA CACACACACA CACACACACA CACACACCCC GTAGTGATAA  15840
AACTATGTAA ATGATATTTC CATAATTAAT ACGTTTATAT TATGTTACTT TTAATGGATG  15900
AATATGTATC GAAGCCCCAT TTCATTTACA TACACGTGTA TGTATATCCT TCCTCCCTTC  15960
CTTCATTCAT TATTTATTAA TAATTTTCGT TTATTTATTT TCTTTTCTTT TGGGGCCGGC  16020
CCGCCTGGTC TTCTGTCTCT GCGCTCTGGT GACCTCAGCC TCCCAAATAG CTGGGACTAC  16080
AGGGATCTCT TAAGCCCGGG AGGAGAGGTT AACGTGGGCT GTGATCGCAC ACTTCCACTC  16140
CAGCTTACGT GGGCTGCGGT GCGGTGGGGT GGGGTGGGGT GGGGTGGGGT GCAGAGAAAA  16200
CGATTGATTG CGATCTCAAT TGCCTTTTAG CTTCATTCAT ACCCTGTTAT TTGCTCGTTT  16260
ATTCTCATGG GTTCTTCTGT GTCATTGTCA CGTTCATCGT TTGCTTGCCT GCTTGCCTGT  16320
TTATTTCCTT CCTTCCTTCC TTCCTTCCTT CCTTCCTTCC TTCCTTCCTT CCCTCCCTTA  16380
CTGGCAGGGT CTTCCTCTGT CTCTGCCGCC CAGGATCACC CCAACCTCAA CGCTTTGGAC  16440
CGACCAAACG GTCGTTCTGC CTCTGATCCC TCCCATCCCC ATTACCTGAG ACTACAGGCG  16500
CGCACCACCA CACCGGCTGA CTTTTATGTT GTTTCTCATG TTTTCCGTAG GTAGGTATGT  16560
GTGTGTGTGT GTGTGTGTGT GTGTGTGTGT GTGTGTGTGT GTGTGTGTGT GTGTGTATCT  16620
ATGTATGTAC GTATGTATGT ATGTATGTGA GTGAGATGGG TTTCGGGGTT CTATCATGTT  16680
GCCCACGCTG GTCTCGAACT CCTGTCCTCA AGCAATCCGC CTGCCTGCCT CGGCCGCCCA  16740
CACTGCTGCT ATTACAGGCG TGAGACGCTG CGCCTGGCTC CTTCTACATT TGCCTGCCTG  16800
CCTGCCTGCC TGCCTGCCTA TCAATCGTCT TCTTTTTAGT ACGGATGTCG TCTCGCTTTA  16860
TTGTCCATGC TCTGGGCACA CGTGGTCTCT TTTCAAACTT CTATGATTAT TATTATTGTA  16920
GGCGTCATCT CACGTGTCGA GGTGATCTCG AACTTTTAGG CTCCAGAGAT CCTCCCGCAT  16980
CGGCCTCCCG GAGTGCTGTG ATGACACGCG TGGGCACGGT ACGCTCTGGT CGTGTTTGTC  17040
GTGGGTCGGT TCTTTCCGTT TTTAATACGG GGACTGCGAA CGAAGAAAAT TTTCAGACGC  17100
ATCTCACCGA TCCGCCTTTT CGTTCTTTCT TTTTATTCTC TTTAGACGGA GTTTCACTCT  17160
TGTCGCCCAG GGTGGAGTAC GATGGCGGCT CTCGGCTCAC CGCACCCTCC GCCTCCCAGG  17220
TTCAAGTGAT TCTCCTGCCT CAGCCTTCCC GAGTAGCTGG AATGACAGAG ATGAGCCATC  17280
GTGCCCGGCT AATTTTTCTA TTTTTAGTAC AGATGGGGTT TCTCCATCTT GGTCAGGCTG  17340
GTCTTCAACT TCCGACCGTT GGAGAATCTT AACTTTCTTG GTGGTGGTTG TTTTCCTTTT  17400
TCTTTTTTTT TCTTTTCTTT TCTTTCCTTC TCCTCCCCCC CCCACCCCCC TTGTCGTCGT  17460
CCTCCTCCTC CTCCTCCTCC TCCTCCTCCT CCTCCTCCTC CTCCTCCTCC TCTTTCATTT  17520
CTTTCAGCTG GGCTCTCCTA CTTGTGTTGC TCTGTTGCTC ACGCTGGTCT CAAACTCCTG  17580
GCCTTGACTC TTCTCCCGTC ACATCCGCCG TCTGGTTGTT GAAATGAGCA TCTCTCGTAA  17640
AATGGAAAAG ATGAAAGAAA TAAACACGAA GACGGAAAGC ACGGTGTGAA CGTTTCTCTT  17700
GCCGTCTCCC GGGGTGTACC TTGGACCCGG AAACACGGAG GGAGCTTGGC TGAGTGGGTT  17760
TTCGGTGCCG AAACCTCCCG AGGGCCTCCT TCCCTCTCCC CCTTGTCCCC GCTTCTCCGC  17820
CAGCCGAGGC TCCCACCGCC GCCCCTGGCA TTTTCCATAG GAGAGGTATG GGAGAGGACT  17880
GACACGCCTT CCAGATCTAT ATCCTGCCGG ACGTCTCTGG CTCGGCGTGC CCCACCGGCT  17940
ACCTGCCACC TTCCAGGGAG CTCTGAGGCG GATGCGACCC CCACCCCCCC GTCACGTCCC  18000
GCTACCCTCC CCCGGCTGGC CTTTGCCGGG CGACCCCAGG GGAACCGCGT TGATGCTGCT  18060
TCGGATCCTC CGGCGAAGAC TTCCACCGGA TGCCCCGGGT GGGCCGGTTG GGATCAGACT  18120
GGACCACCCC GGACCGTGCT GTTCTTGGGG GTGGGTTGAC GTACAGGGTG GACTGGCAGC  18180
CCCAGCATTG TAAAGGGTGC GTGGGTATGG AAATGTCACC TAGGATGCCC TCCTTCCCTT  18240
CGGTCTGCCT TCAGCTGCCT CAGGCGTGAA GACAACTTCC CATCGGAACC TCTTCTCTTC  18300
CCTTTCTCCA GCACACAGAT GAGACGCACG AGAGGGAGAA ACAGCTCAAT AGATACCGCT  18360
GACCTTCATT TGTGGAATCC TCAGTCATCG ACACACAAGA CAGGTGACTA GGCAGGGACA  18420
CAGATCAAAC ACTATTTCCG GGTCCTCGTG GTGGGATTGG TCTCTCTCTC TCTCTCTCTC  18480
TCTCTCTCTC TCTCTCTCTC TCTCGCACGC GCACGCGCGC ACACACACAC ACAATTTCCA  18540
TATCTAGTTC ACAGAGCACA CTCACTTCCC CTTTTCACAG TACGCAGGCT GAGTAAAACG  18600
CGCCCCACCC TCCACCCGTT GGCTGACGAA ACCCCTTCTC TACAATTGAT GAAAAAGATG  18660
ATCTGGGCCG GGCACGCTAG CTCACGCCTG TCACTCCGGC ACTTTGGGAG GCCGAGGCGG  18720
GTGGATCGCT TGGGGCCGGG AGTTCGAGAC CAGGCTGGCC GACGTGGCGA AACCCCGTCT  18780
CTCTGAAAAA TAGAACGATT AGCCGGGCCT GGTGGCGTGG GCTTGGAATC ACGACCGCTC  18840
GGGAGACTGG GGCGGGCGAC TTGTTCCAAC CGGGGAGGCC GAGGCCGCGA TGAGCTGAGA  18900
TCGTGCCGTG GCGATGCGGC CTGGATGACG GAGCGAGACC CCGTCTCGAG AGAATCATGA  18960
TGTTATTATA AGATGAGTTG TGCGCGGTGA TGGCCGCCTG TAGTCGCGGC TACTCGGGAG  19020
```

```
GCTGAGACGA GGAGAAGATC ACTTGAGGCC CCACAGGTCG AGGCTTCGGT CGGCCGTGAC  19080
CCACTGTATC CTGGGCAGTC ACCGGTCAAG GAGATATGCC CCTTCCCCGT TTGCTTTTCT  19140
TTTCTTCCCT TCTCTTTTCT TCTTTTTGCT TCTCTTTTCT TTCTTTCTTT CTTTCTTTCT  19200
TTCTTTCTTT CTTTCTTTCT TTTTCTTTTT CTCTCTTCCC CTCTTTCTTT CCTGCCTTCC  19260
TGCCTTTCTT CTTTTCTTCT TTCCTCCCTT CCTCCCTTCC TTCTTTCCTC CCGCCTCAGC  19320
CTCCCAAAGT GCTGGGATGA CTGGCGGGAG GCACCATGCC TGCTTGGCCC AAAGAGACCC  19380
TCTTGGAAAG TGAGACGCAG AGAGCGCCTT CCAGTGATCT CATTGACTGA TTTAGAGACG  19440
GCATCTCGCT CCGTCACCCC GGCAGTGGTG CCGTCGTAAC TCACTCCCTG CAGCGTGGAC  19500
GCTCCTGGAC TCGAGCGATC CTTCCACCTC AGCCTCCAGA GTACAGAGCC TGGGACCGCG  19560
GGCACGCGCC ACTGTGCCCA CACCGTTTTT AATTGTTTTT TTTTCCCCCG AGACAGAGTT  19620
TCACTCTCGT GGCCTAGACT GCAGTGCGGT GGCGCGATCT TGGCTCACCG CAACCTCTGC  19680
CTCCCGGTTT CAAGCGATTC TCCTGCATCG GCCTCCTGAG TAGCCGGGAT TGCGGGCATG  19740
CGCTGCCACG TCTGGCTGAT TTCGTATTTT TAGTGGAGAC GGGGCTTCTC CATGTCGATC  19800
GGGCTGGTTT CGAACTCCCG ACCTCAGGTG ATCCGCCCTC CCCGGCCTCC GGAAGTGCTG  19860
GGATGACAGG CGTGAGCCAC CGCGCCCGGC CTTCATTTTT AAATGTTTTC CCACAGACGG  19920
GGTCTCATCA TTTCTTTGCA ACCCTCCTGC CCGGCGTCTC AAAGTGCTGG CGTGACGGGC  19980
GTGAGCCACT GCGCCTGGAC TCCGGGGAAT GACTCACGAC CACCATCGCT CTACTGATCC  20040
TTTCTTTCTT TCTTTCTTTC TTTCTTTCTT TCTTTCTTTC TTTCTTTCTT TCTTTCTTGA  20100
TGAATTATCT TATGATTTAT TTGTGTACTT ATTTTCAGAC GGAGTCTCGC TCTGGGCGGG  20160
GCGAGGCGAG GCGAGGCACA GCGCATCGCT TTGGAAGCCG CGGCAACGCC TTTCAAAGCC  20220
CCATTCGTAT GCACAGAGCC TTATTCCCTT CCTGGAGTTG GAGCTGATGC CTTCCGTAGC  20280
CTTGGGCTTC TCTCCATTCG GAAGCTTGAC AGGCGCAGGG CCACCCAGAG GCTGGCTGCG  20340
GCTGAGGATT AGGGGGTGTG TTGGGGCTGA AAACTGGGTC CCCTATTTTT GATACCTCAG  20400
CCGACACATC CCCCGACCGC CATCGCTTGC TCGCCCTCTG AGATCCCCCG CCTCCACCGC  20460
CTTGCAGGCT CACCTCTTAC TTTCATTTCT TCCTTTCTTG CGTTTGAGGA GGGGGTGCGG  20520
GAATGAGGGT GTGTGTGGGG AGGGGGTGCG GGGTGGGGAC GGAGGGGAGC GTCCTAAGGG  20580
TCGATTTAGT GTCATGCCTC TTTCACCACC ACCACCACCA CCGAAGATGA CAGCAAGGAT  20640
CGGCTAAATA CCGCGTGTTC TCATCTAGAA GTGGGAACTT ACAGATGACA GTTCTTGCAT  20700
GGGCAGAACG AGGGGGACCG GGGACGCGGA AGTCTGCTTG AGGGAGGAGG GGTGGAAGGA  20760
GAGACAGCTT CAGGAAGAAA ACAAAACACG AATACTGTCG GACACAGCAC TGACTACCCG  20820
GGTGATGAAA TCATCTGCAC ACTGAACACC CCCGTCACAA GTTTACCTAT GTCACAATCT  20880
TGCACATGTA TCGCTTGAAC GACAAATAAA AGTTAGGGGG GAGAAGAGAG GAGAGAGAGA  20940
GAGAGAGAGA GACAGAGAGA GACAGAGAGA GAGAGAGAGG AGGGAGAGAG GAAAACGAAA  21000
CACCACCTCC TTGACCTGAG TCAGGGGGTT TCTGGCCTTT TGGGAGAACG TTCAGCGACA  21060
ATGCAGTATT TGGGCCCGTT CTTTTTTTTT CTTCTTCTTT TCTTTCTTTT TTTTTGGACT  21120
GAGTCTCTCT CGCTCTGTCA CCCAGGCTGC GGTCGCGGTG GCGCTCTCTC GGCTCACTGA  21180
AACCTCTGCT TCCCGGGTTC CAGTGATTCT TCTTCGGTAG CTGGGATTAC AGGCGCACAC  21240
CATGACGGCG GGCTCATATT CCTATTTTCA GTAGAGACGG GGTTTCTCCA CGTTGGCCAC  21300
GCTGGTCTCG AACTCCTGAC CTCAAATGAT CCGCCTTCCT GGGCCTCCCA AAGTGCTGGA  21360
AACGACAGGC CTGAGCCGCC GGGATTTCAG CCTTTAAAAG CGCGGCCCTG CCACCTTTCG  21420
CTGTGGCCCT TACGCTCAGA ATGACGTGTC CTCTCTGCCG TAGGTTGACT CCTTGAGTCC  21480
CCTAGGCCAT TGCACTGTAG CCTGGGCAGC AAGAGCCAAA CTCCGNNCCC CCACCTCCTC  21540
GCGCACATAA TAACTAACTA ACAAACTAAC TAACTAACTA AACTAACTAA CTAACTAAAA  21600
TCTCTACACG TCACCCATAA GTGTGTGTTC CCGTGAGAGT GATTTCTAAG AAATGGTACT  21660
GTACACTGAA CGCAGTGGCT CACGTCTGTC ATCCCGAGGT CAGGAGTTCG AGACCAGCCC  21720
GGCCAACGTG GTGAAACCCC GTCTCTACTG AAAATACGAA ATGGAGTCAG GCGCCGTGGG  21780
GCAGGCACCT GTAACCCCAG CTACTCGGGA GGCTGGGGTG GAAGAATTGC TTGAACCTGG  21840
CAGGCGGAGG CTGCAGTGAC CCAAGATCGC ACCACTGCAC TACAGCCTGG GCGACAGAGT  21900
GAGACCCGGT CTCCAGATAA ATACGTACAT AAATAAATAC ACACATACAT ACATACATAC  21960
ATACATACAT ACATACATAC ATCCATGCAT ACAGATATAC AAGAAAGAAA AAAAGAAAAG  22020
AAAAGAAAGA GAAAATGAAA GAAAAGGCAC TGTATTGCTA CTGGGCTAGG GCCTTCTCTC  22080
TGTCTGTTTC TCTCTGTTCG TCTCTGTCTT TCTCTCTGTG TCTCTTTCTC TGTCTGTCTG  22140
TCTCTTTCTT TCTCTCTGTC TCTGTCTCTG TCTTTGTCTC TCTCTCTCCC TCTCTGCCTG  22200
TCTCACTGTG TCTGTCTTCT GTCTTACTCT CTTTCTCTCC CCGTCTGTCT CTCTCTCTCT  22260
CTCTCCCTCC CTGTTTGTTT CTCTCTCTCC CTCCCTGTCT GTTTCTCTCT CTCTCTTTCT  22320
GTCTGTTTCT GTCTCTCTCT GTCTGTCTAT GTCTTTCTCT GTCTGTCTCT TTCTCTGTCT  22380
GTCTGCCTCT CTCTTTCTTT TTCTGTGTCT CTCTGTCGGT CTCTCTCTCT CTGTCTGTCT  22440
GTCTGTCTCT CTCTCTCTCT CTCTGTGCCT ATCTTCTGTC TTACTCTCTT TCTCTGCCTG  22500
TCTGTCTGTC TCTCCCTCCC TTTCTGTTTC TCTCTCTCTC TCTCTCTCTC TCCCCCTCTC  22560
CCTGTCTGTT TCTCTCCGTC TCTCTCTCTT TCTGTCTGTT TCTCACTGTC TCTCTCTGTC  22620
CATCTCTCTC TCTCTCTGTC TGTCTCTTTC GTTCTCTCTG TCTGTCTGTC TCTCTCTCTC  22680
TCTCTCTCTC TCTCTCTCTC TCCCTGTCTG TCTGTTTCTC TCTATCTCTC GCTGTCCATC  22740
TCTGTCTTTC TATGTCTGTC TCTTTCTCTG TCAGTCTGTC AGACACCCCC GTGCCGGGTA  22800
GGGCCCTGCC CCTTCCACGA AAGTGAGAAG CGCGTGCTTC GGTGCTTAGA GAGGCCGAGA  22860
GGAATCTAGA CAGGCGGGCC TTGCTGGGCT TCCCCACTCG GTGTATGATT TCGGGAGGTC  22920
```

```
GAGGCCGGGT CCCCGCTTGG ATGCGAGGGG CATTTTCAGA CTTTTCTCTC GGTCACGTGT  22980
GGCGTCCGTA CTTCTCCTAT TTCCCCGATA AGCTCCTCGA CTTCAACATA AACGGCGTCC  23040
TAAGGGTCGA TTTAGTGTCA TGCCTCTTTC ACCGCCACCA CCGAAGATGA AAGCAAAGAT  23100
CGGCTAAATA CCGCGTGTTC TCATCTAGAA GTGGGAACTT ACAGATGACA GTTCTTGCAT  23160
GGGCAGAACG AGGGGGACCG GGNACGCGGA AGCCTGCTTG AGGGRGGAGG GGYGGAAGGA  23220
GAGACAGCTT CAGGAAGAAA ACAAAACACG AATACTGTCG GACACAGCAC TGACTACCCG  23280
GGTGATGAAA TCATCTGCAC ACTGAACACC CCCGTCACAA GTTTACCTAT GTCACAGTCT  23340
TGCTCATGTA TGCTTGAACG ACAAATAAAA GTTCGGGGGG GAGAAGAGAG GAGAGAGAGA  23400
GAGAGACGGG GAGAGAGGGG GGAGAGGGGG GGGGAGAGAG AGAGAGAGAG AGAGAGAGAG  23460
AGAGAGAGAG AGAAAGAGAA GTAAAACCAA CCACCACCTC CTTGACCTGA GTCAGGGGGT  23520
TTCTGGCCTT TTGGGAGAAC GTTCAGCGAC AATGCAGTAT TTGGGCCCGT TCTTTTTTTC  23580
TTCTTCTTCT TTTCTTTCTT TTTTTTTGGA CTGAGTCTCT CTCGCTCTGT CACCCAGGCT  23640
GCGGTGCGGT GGCGCTCTCT CGGCTCACTG AAACCTCTGC TTCCCGGGTT CCAGTGATTC  23700
TTCTTCGGTA GCTGGGATTA CAGGTGCGCA CCATGACGGC CGGCTCATCG TTCTATTTTT  23760
AGTAGAGACG GGGTTTCTCC ACGTTGGCCA CGCTGGTCTC GAACTCCTGA CCACAAATGA  23820
TCCACCTTCC TGGGCCTCCC AAAGTGCTGG AAACGACAGG CCTGAGCCGC CGGGATTTCA  23880
GCCTTTAAAA GCGCGCGGCC CTGCCACCTT TCGCTGCGGC CCTTACGCTC AGAATGACGT  23940
GTCCTCTCTG CCATAGGTTG ACTCCTTGAG TCCCCTAGGC CATTGCACTG TAGCCTGGGC  24000
AGCAAGAGCC AAACTCCGTC CCCCCACCTC CCCGCGCACA TAATAACTAA CTAACTAACT  24060
AACTAACTAA AATCTCTACA CGTCACCCAT AAGTGTGTGT TCCCGTGAGG AGTGATTTCT  24120
AAGAAATGGT ACTGTACACT GAACGCAGGC TTCACGTCTG TCATCCCGAG GTCAGGAGTT  24180
CGAGACCAGC CCGGCCCACG TGGTGAAACC CCCGTCTCTA CTGAAAATAC GAAATGGAGT  24240
CAGGCGCCGT GGGGCAGGCA CCTGTAACCC CAGCTACTCG GGAGGCTGGG GTGGAAGAAT  24300
TGCTTGAACC TGGCAGGCGG AGGCTGCAGT GACCCAAGAT CGCACCACTG CACTACAGCC  24360
TGGGCGACAG AGTGAGACCC GGTCTCCAGA TAAATACGTA CATAAATAAA TACACACATA  24420
CATACATACA TACATACAAC ATACATACAT ACAGATATAC AAGAAAGAAA AAAAGAAAAG  24480
AAAAGAAAGA GAAAATGAAA GAAAAGGCAC TGTATTGCTA CTGGGCTAGG GCCTTCTCTC  24540
TGTCTGTTTC TCTCTGTTCG TCTCTGTCTT TCTCTCTGTG TCTCTTTCTC TGTCTGTCTG  24600
TCTGTCTGTC TGTCTGTCTC TTTCTTTCTT TCTGTCTCTG TCTTTGTCCC TCTCTCTCCC  24660
TCTCTGCCCT GTCTCACTGT GTCTGTCTTC TATCTTACTC TCTTTCTCTC CCCGTCTGTC  24720
TCTCTCTCAC TCCCTCCCTG TCTGTTTCTC TCTCTCTCTC TTTCTGTCTG TTTCTGTCTC  24780
TCTCTGTCTG CCTCTCTCTT TCTCTATCTG TCTCTTTCTC TGTCTGTCTG CCCCTCTCTT  24840
TCTTTTTCTG TGTCTCTCTG TCTGTCTCTC TCTCTCTCTG TGCCTATCTT CTGTCTTACT  24900
CTCTTTCTCT GCCTGTCTGT CTGTCTCTCT CTGTCTCTCC CTCCCTTTCT GCTTCTCTCT  24960
CTCTCTCTCT CTCTNNNCCC TCCCTGTCTG TTTCTCTCTG TCTCCCTCTC TTTCTGTCTG  25020
TTTCTCACTG TCTCTCTCTG TCTGTCTGTT TCATTCTCTC TGTCTCTGTC TCTGTCTCTC  25080
TCTCTCTCTG TCTCTCCCTC TCTGTGTGTA TCTTTTGTCT TACTCTCCTT CTCTGCCTGT  25140
CCGTCTGTCT GTCTGTCTCT CTCTCTCCCT GTCCCTCTCT CTTTCTGTCT GTTTCTCTCT  25200
CTCTCTCTCT CTCTCTCTCT CTGTCTCTGT CTTTCTCTGT CTGTCCCTTT CTCTGTCTGT  25260
CTGCCTCTCT CTTTCTCTTT CTGTGTCTCT CTGTCTCTCT CTCTGTGCCT ATCTTCTGTC  25320
TTACTCTCTT TCTCTGCCTG TCTATCTGTC TGTCTCTCTC TGTCTCTCTC CCTGCCTTTC  25380
TGTTTCTCTC TCTCTCCCTC TCTCGCTCTC TCTGTCTTTC TCTCTTTCTC TCTGTTTCTC  25440
TGTCTCTCTC TGTCCGTCTC TGTCTTTTTC TGTCTGTCTG TCTCTCTCTT TCTTTCTGTC  25500
GTCTGTCTCT GTCTCTGTCT CTGTCTCTCT CTCTCTCTCT CTCCTTGTCT CTCTCACTGT  25560
GTCTGTCTTC TGTCTTACTC TCCTTCTCTG CCTGTCCATC TGTCTGTCTG TCTCTCTCTC  25620
TCTCTCCCTA CCTTTCTGTT TCTCTCTCGC TAGCTCTCTC TCTCTCTGCC TGTTTCTCTC  25680
TTTCTCTCTC TGTCTTTCTC TGTCTGTCTC TTTCTCTGTC TGTCTGTCTC TTTCTCTCTG  25740
TCTCTGTCTC TGTCTCTCTC TCTCTCTCTC TCTCTCTCTC TGCCTCTCTC ACTGTGTCTG  25800
TCTTCTGTCT TATTCTCTTT CTCTCTCTGT CTCTCTCTCT CTCTCCTTTA CTGTCTGTTT  25860
CTCTCTCTCT CTCTCTCTTT CTGCCTGTTT CTCTCTGTCT GTCTCTGTCT TTCTCTGTCT  25920
GTCTGCCTCT CTCTTTCTTT TTCTGCGTCT CTCTGTCTCT CTCTCTCTCT CTCTGTTCCT  25980
ATCTTCTGTC TTACTCTGTT TCCTTGCCTG CCTGCCTGTC TGTGTGTCTG TCTCTCTCTC  26040
TCTCTCTCTC TCTCTCTCCC TCCCTTTCTC TTTCTCTGTC TCTCTCTCTC TTTCTGGGTG  26100
TTTCTCTCTG TCTCTCTGTC CATCTCTGTC TTTCTATGTC TGTCTCTCTC TTTCTCTCTG  26160
TCTCTGTCTC TGCCTCTCTC TCTCTCTCTC TCTCTCTCTC TCTGTCTGTC TCTCTCACTG  26220
TGTGTGTCTG TCTTCTGTCT TACTCTCCTT CTCTGCCTGT CCGTCTGTCT GTCTGTCTCT  26280
CCCTCTCTCT CCCTCCCTTT CTGTTTCTCT CTCTCTCTCT TTCTGTCTGT TTCTCTCTTT  26340
CTCTCTCTGT CTGTCTCTTT CTCTGTCTGT CTGTCTCTCT CTTTCTTTTT CTCTGTCTCT  26400
CTGTCTCTCT CTGTGTCTGT CTCTCTGTCT GTGCCTATCT TCTGTCTTAC TCTCTTTCTC  26460
TGGCTGTCTG CCTGTCTCTC TCTCTCTCTC TGTCTGTCTC CGTCCCTCTC TCCCTGTCTG  26520
TCTGTTTCTC TCTCTGCCTC TCTCTCTCTC TGTCTGTCTC TTTCTCTGTC TGTCTGTCTC  26580
TCTCTTTCTT TTTCTCTGTC TCTCTGTCTC TCTCTGTGTC TGTCTCTCTT TCTGTGCCTA  26640
TCTTCTGTCT TACTCTCTTT CTCTGGCTGT CTGCCTGTCT CTCTCTCTCT GCCTGTCTCC  26700
GTCCCTCCCT CCCTGTCTGT CTGTTTCTCT CTCTGTCTCT GTCTCTCTGT CCATCTCTGT  26760
CTGTCTCTTT CTCTTTCTCT CTCTCTGTCT CTGTCTCTCT CTCTCTCTGC CTGTCTCTCT  26820
```

```
CACTGTGTCT GTCTTCTGTC TTACTCTCTT TCTCTTGCCT GCCTCTCTGT CTGTCTGTCT  26880
CTCTCCCTCC ATGTCTCTCT CTCTCTCTCA CTCACTCTCT CTCCGTCTCT CTCTCTTTCT  26940
GTCTGTTTCT CTCTCTGTCT GTCTCTCTCC CTCCATGTCT CTCTCTCTCT CTCTCACTCA  27000
CTCTCTCTCC GTCTCTCTCT CTCTTTCTGT CTGTTTCTCT CTCTGTCTGT CTCTCTCCCT  27060
CCATGTCTCT CTCTCTCCCT CTCACTCACT CTCTCTCCGT CTCTCTCTCT CTTTCTGTCT  27120
GTTTCTTTGT CTGTCTGTCT GTCTGTCTGT CTGTCTCTCT CTCTCTCTCT CTCTCTCTCT  27180
CTCTCTGTTT GTCTTTCTCC CTCCCTGTCT GTCTGTCTGT CTCTCTCTCT CTGTCTCTGT  27240
CTCTGTCTCT CTCTCTTTCT CTTTCTGTCT GTTTCTCTCT ATCTCTCGCT GTCCATCTCT  27300
GTCTTTCTAT GTCTGTCTCT TTCTCTGTCA GTCTGTCAGA CACACCCGTG CCGGTAGGGC  27360
CCTGCCCTTC CACGAGAGTG AGAAGCGCGT GCTTCGGTGC TTAGAGAGGC CGAGAGGAAT  27420
CTAGACAGGC GGGCCTTGCT GGGCTTCCCC ACTCGGTGTA CGATTTCGGG AGGTCGAGGC  27480
CGGGTCCCCG CTTGGATGCG AGGGGCATTT TCAGACTTTT CTCTCGGTCA CGTGTGGCGT  27540
CCGTACTTCT CCTATTTCCC CGATAAGTCT CCTCGACTTC AACATAAACT GTTAAGGCCG  27600
GACGCCAACA CGGCGAAACC CCGTCTCTAC TAAAAATACA AAGCTGAGTC GGGAGCGGTG  27660
GGGCAGGCCC TGTAATGCCA GCTCCTCGGG AGGCTGAGGC GGGAGAATCG CTTGAACCAG  27720
GGAAGCGGAG GCTGCAGGGA GCCGAGATCG CGCCACTGCA CTACGGCCCA GGCTGTAGAG  27780
TGAGTGAGAC TCGGTCTCTA AATAAATACG GAAATTAATT AATTCATTAA TTCTTTTCCC  27840
TGCTGACGGA CATTTGCAGG CAGGCATCGG TTGTCTTCGG GCATCACCTA GCGGCCACTG  27900
TTATTGAAAG TCGACGTTGA CACGGAGGGA GGTCTCGCCG ACTTCACCGA GCCTGGGGCA  27960
ACGGGTTTCT CTCTCTCCCT TCTGGAGGCC CCTCCCTCTC TCCCTCGTTG CCTAGGGAAC  28020
CTCGCCTAGG GAACCTCCGC CCTGGGGGCC CTATTGTTCT TTGATCGGCG CTTTACTTTT  28080
CTTTGTGTTT TGGCGCCTAG ACTCTTCTAC TTGGGCTTTG GGAAGGGTCA GTTTAATTTT  28140
CAAGTTGCCC CCCGGCTCCC CCCACTACCC ACGTCCCTTC ACCTTAATTT AGTGAGNCGG  28200
TTAGGTGGGT TTCCCCCAAA CCGCCCCCCC CCCCCCGCCT CCCAACACCC TGCTTGGAAA  28260
CCTTCCAGAG CCACCCCGGT GTGCCTCCGT CTTCTCTCCC CTTCCCCCAC CCCTTGCCGG  28320
CGATCTCATT CTTGCCAGGC TGACATTTGC ATCGGTGGGC GTCAGGCCTC ACTCGGGGGC  28380
CACCGTTTTT GAAGATGGGG GCGGCACGGT CCCACTTCCC CGGAGGCAGC TTGGGCCGAT  28440
GGCATAGCCC CTTGACCCGC GTGGGCAAGC GGGCGGGTCT GCAGTTGTGA GGCTTTTCCC  28500
CCCGCTGCTT CCCGCTCAGG CCTCCCTCCC TAGGAAAGCT TCACCCTGGC TGGGTCTCGG  28560
TCACCTTTTA TCACGATGTT TTAGTTTCTC CGCCCTCCGG CCAGCAGAGT TTCACAATGC  28620
GAAGGGCGCC ACGGCTCTAG TCTGGGCCTT CTCAGTACTT GCCCAAAATA GAAACGCTTT  28680
CTGAAAACTA ATAACTTTNC TCACTTAAGA TTTCCAGGGA CGGCGCCTTG GCCCGTGTTT  28740
GTTGGCTTGT TTTGTTTCGT TCTGTTTTGT TTTGTTCGTG TTTTTCCTTT CTCGTATGTC  28800
TTTCTTTTCA GGTGAAGTAG AAATCCCCAG TTTTCAGGAA GACGTCTATT TTCCCCAAGA  28860
CACGTTAGCT GCCGTTTTTT CCTGTTGTGA ACTAGCGCTT TTGTGACTCT CTCAACGCTG  28920
CAGTGAGAGC CGGTTGATGT TTACNATCCT TCATCATGAC ATCTTATTTT CTAGAAATCC  28980
GTAGGCGAAT GCTGCTGCTG CTCTTGTTGC TGTTGTTGTT GTTGTTGTTG TCGTCGTTGC  29040
TGTTGTCGTT GTCGTTGTTG TTGTCGTTGT CGTTGTTTTC AAAGTATACC CCGGCCACCG  29100
TTTATGGGAT CAAAAGCATT ATAAAATATG TGTGATTATT TCTTGAGCAC GCCCTTCCTC  29160
CCCCTCTCTC TGTCTCTCTG TCTGTCTCTG TCTCTCTCTT TCTCTGTCTG TCTTCTCTCT  29220
CTCTCTCTCT CTGTGTCTCT CTCTCTCTGC CTGTCTGTTT CTCTCTCTCT GCCTCTCTCT  29280
CTCTCTCTCT CTCTGCCTGT CTCTCTCACT GTGTCTGTCT TCTGTCTTAC TCCCTTTCTC  29340
TGTCTGTCTG TCGGTCTCTC TCTCTCTCTC TCCCTGTCTG TATGTTTCTC TCTGTCTCTG  29400
TCTCTCTCTC TCTTTCTGTT TCTCTCTCTC CGTCTCTGTC TTTCTCTGAC TGTCTCTCTC  29460
TTTCCTTCTC TCTGTCTCTC TCTGCCTGTC TCTCTCACTC TGTCTTCTGT CTTATCTCTC  29520
TCTCTGCCTG CCTGTCTCTC TCACTCTCTC TCTCTGTGTG TCTCTCTCTC TCTTTCTGTT  29580
TCTCTCTGTC TCTCTGTCCG TCTCTGTCTT TCTCTGTCTG TCTCTTTGTC TGTCTGTCTT  29640
TGTCTTTCCT TCTCTCTGTC TCTGTCTCTC TCACTGTGTC TGTCTTCTGT CTTAGTCTCT  29700
CTCTCTCTCT CTCCCTGTCT GTCTGTCTCT CTCTCTCTCT CCCCCTGTCT GTTTCTCTCT  29760
CTCTCTCTCT CTCTCTCTCT CTCTGTCTTT GTCTTTCTTT CTGTCTCTGT CTCTCTCTCT  29820
CTCTCTGTGT GTCTGTCTTC TGTCTTACTG TCTTTCTCTG CCTGTCTGTC TGTCTGTCTC  29880
TCTCTGTCTG TCTCTCTCTC TCTCTCCCCC TGTCGGCTGT TTCTCTGTCT CTGTCTGTGT  29940
CTCTCTTTCT GTCTGTTTCT CTCTGTCTGT CTTTCTCTCT CTGTCTCTTT CTCTCTGTCT  30000
CTCTGTCTGT CTCTGTCTCT CTCTCTGTCT CTCTCTCTCT GTGGGGGTGT GTGTGTGTGT  30060
GTGTATGTGT GTGTGTGTGT GTGTGTGTGT CTGCCTTCTG TCTTACTCTC TTTCTCTGCC  30120
TGTCTGTCTG CCTGTCTGTT TGTCTCTCTC TCTCTGCCTG TCTCTCTCCC TTCCTGTCTG  30180
TTTCTCTCTC TTTCTGTTTC TCTCTGTCTC TGTCCATCTC TGTCTTTCTC CGTCTGTCTC  30240
TTTATCTGTC TCTCTCCGTC TGTCTCTTTA TCTGTCTCTC TCTCTCTTTC TGTCTTTCTC  30300
TCTCTGTGTA TCGTTGTCTC TCTCTGTCTG TCTCTGTCTC TGTCTCTCTG TCTCTCTCTC  30360
TCTCTCTCTC TCTCTGTCTG TCTGTCCGTC TGTCTGTCTC GGTCTCTGCG TCTCGCTATC  30420
TCCCGCCCTC TCTTTTTTTG CAAAAGAAGC TCAAGTACAT CTAATCTAAT CCCTTACCAA  30480
GGCCTGAATT CTTCACTTCT GACATCCCAG ATTTGATCTC CCTACAGAAT GCTGTACAGA  30540
ACTGGCGAGT TGATTTCTGG ACTTGGATAC CTCATAGAAA CTACATATGA ATAAAGATCC  30600
AATCCTAAAA TCTGGGGTGG CTTCTCCCTC GACTGTCTCG AAAAATCGTA CCTCTGTTCC  30660
CCTAGGATGC CGGAAGAGTT TTCTCAATGT GCATCTGCCC GTGTCCTAAG TGATCTGTGA  30720
```

```
CCGAGCCCTG TCCGTCCTGT CTCAAATATG TACGTGCAAA CACTTCTCTC CATTTCCACA  30780
ACTACCCACG GCCCCTTGTG GAACCACTGG CTCTTTGAAA AAAATCCCAG AAGTGGTTTT  30840
GGCTTTTTGG CTAGGAGGCC TAAGCCTGCT GAGAACTTTC CTGCCCAGGA TCCTCGGGAC  30900
CATGCTTGCT AGCGCTGGAT GAGTCTCTGG AAGGACGCAC GGGACTCCGC AAAGCTGACC  30960
TGTCCCACCG AGGTCAAATG GATACCTCTG CATTGGCCCG AGGCCTCCGA AGTACATCAC  31020
CGTCACCAAC CGTCACCGTC AGCATCCTTG TGAGCCTGCC CAAGGCCCCG CCTCCGGGGA  31080
GACTCTTGGG AGCCCGGCCT TCGTCGGCTA AAGTCCAAAG GGATGGTGAC TTCCACCCAC  31140
AAGGTCCCAC TGAACGGCGA AGATGTGGAG CGTAGGTCAG AGAGGGGACC AGGAGGGGAG  31200
ACGTCCCGAC AGGCGACGAG TTCCCAAGGC TCTGGCCACC CCACCCACGC CCCACGCCCC  31260
ACGTCCCGGG CACCCGCGGG ACACCGCCGC TTTATCCCCT CCTCTGTCCA CAGCCGGCCC  31320
CACCCCACCA CGCAACCCAC GCACACACGC TGGAGGTTCC AAAACCACAC GGTGTGACTA  31380
GAGCCTGACG GAGCGAGAGC CCATTTCACG AGGTGGGAGG GGTGGGGGTG GGGTGGGTTG  31440
GGGGTTGTGG GGTCTGTGGC GAGCCCGATT CTCCCTCTTG GGTGGCTACA GGCTAGAAAT  31500
GAATATCGCT TCTTGGGGGG AGGGGCTTCC TTAGGCCATC ACCGCTTGCG GGACTACCTC  31560
TCAAACCCTC CCTTGAGGCC ACAAAATAGA TTCCACCCCA CCCATCGACG TTTCCCCCGG  31620
GTGCTGGATG TATCCTGTCA AGAGACCTGA GCCTGACACC GTCGAATTAA ACACCTTGAC  31680
TGGCTTTGTG TGTTTGTTTG TTTCTGAGAT GGAGTCTTGC TCTGTCCCCC AGGCTGGAGT  31740
GCAGTGGCGT GATCTCAGCT CACTGGAACC TCTGCCTCCT GGGTTCAAGT GATTCTCCTG  31800
TCTCAGCGCC ACCATGGCCG GCTCATTTTT TTTTTTTTTT TTTTTGGTAG ACACGGGGTT  31860
TCACCCTCTT TCATTGGTTT TCACTGGAGA TTCTAGATTC GAGCCACACC TCATTCCGTG  31920
CCACAGAGAG ACTTCTTTTT TTTTTTTTTT TTTTTAAGCG CAACGCAACA TGTCTGCCTT  31980
ATTTGAGTGG CTTCCTATAT CATTATAATT GTGTTATAGA TGAAGAAACG GTATTAAACA  32040
CTGTGCTAAT GATAGTGAAA GTGAAGACAA AAGAAAGGCT ATCTATTTTG TGGTTAGAAT  32100
AAAGTTGCTC AGTATTTAGA AGCTACCTAA ATACGTCAGC ATTTACACTC TTCCTAGTAA  32160
AAGCTGGCCG ATCTGAATAA TCCTCCTTTA AACAAACACA ATTTTTGATA GGGTTAAGAT  32220
TTTTTTAAGA ATGCGACTCC TGCAAAATAG CTGAACAGAC GATACACATT TAAAAAAATA  32280
ACAACACAAG GATCAACCAG ACTTGGGAAA AAATCGAAAA CCACACAAGT CTTATGAAGA  32340
ACTGAGTTCT TAAAATAGGA CGGAGAACGT AGCTATCGGA AGAGAAGGCA GTATTGGCAA  32400
GTTGATTGTT ACGTTGGTCA GCAGTAGCTG GCACTATCTT TTTGGCCATC TTTCGGGCAA  32460
TGTAACTACT ACAGCAAAAT GAGATATGAT CCATTAAACA ACATATTCGC AAATCAAAAA  32520
GTGTTTCAGT AATATAATGC TTCAGATTTA GAAGCAAATC AAATGATAGA ACTCCACTGC  32580
TGTAATAAGT CACCCCAAAG ATCACCGTAT CTGACAAAAT AACTACCACA GGGTTATGAC  32640
TTCAGAATCA TACTTTCTTC TTGATATTTA CTTATGTATT TATTTTTTTT AATTTATTTC  32700
TCTTGAGACG CGTCTCGCTC TGTCGCCCAG GCTGGAGTGC GATGGTGTGA TCTCGGCTCA  32760
CTGCAACCGC CACCTCCCTG GGTTCAAGCG ATTCTCCTGC CTCAGCCTCC CGAGTAGCTG  32820
GGACTACAGG TGCCCGCCAC CACGCCCAGC TAATCTTTAT ACTTTTAATA GAGACGGGGT  32880
TTCACCGTGT CGGCCCGGAT GGTCTCGATC TCTTGACCTC GTGACCCGCC CGCCTCGGCC  32940
TCCCAAAGTG CTGGGATGAC AGGCGTGAGC CACTGAGCCC GGCCTTCTCT TGACGTTTAA  33000
ACTATGAAGT CAGTCCAGAG AAACGCAATA AATGTCAACG GTGAGGATGG TGTTGAGGCA  33060
GAAGTAGGAC CACACTTTTT CCTATCTTAT TCAGTTGATA ACAATATGAC CTAGGTAGTA  33120
ATTTCCTATG TGCCTACTTA TACACGAGTA CAAAAGAGTA AAACAGAGAG ACTGCTAAAT  33180
TAAAGGGTAC GTGAAGTTCT TCATAGTAAC TCCGTAAACT GGAACACTGT CAAAAAGCAG  33240
CAGCTAGTGA ATTGTTTCCA TGTATTTTTC TATTATCCAA TAAGTGAACT ATGCTATTCC  33300
TTTCCAGTCT CCCAAGCACT TCTTGTCCCC ATCACCACTT CGGTGCTCGA AGAAAAAGTA  33360
AGCAAATCAA GGAACACAAG CTAAAGAAAC ACACACACAA ACCAAAGACA ACTACAGCGT  33420
CTGCAAAAGT TTGCTAGAAG ACTGAAACTG TTGAGTATAA GGATCTGGTA TTCTACGATC  33480
ATGAGTTCAC TTCAGAGTTT GTTCAAGACA TACGTTTCGT AAGGAAACAT CTTAGTTAGA  33540
AGTTATTCAG CAGTAGGTAC CATCCCTAAG TATTTTTCAC CAAATCCGTG ACAATAAAGA  33600
GCTATCTAAC CAGAAAAATT AGCGAGTACG GGCACCATCC ATAGGGCTTT GTCTTTACGC  33660
TTCATTAGCA CTTACCATGC CTTACAATGT CTAGGATTGA CCCTGATAGC ATTTCGAAAA  33720
CAAGCTAATG CTTTGTCCAG TTCTTCAGTG AAGACAACTC ACGCCCTAAT GCGCTATAGG  33780
CATAAGCATC ATTTGGATCC ACTTCGAGAG TTCTCTGGAA GAATTGAATC GCAATATCGT  33840
GTTCCCGTTT GCAGACCGAA ACAGTTTCCC TGCAGCACAC CAGGCCTCTG GCTGGCGAAT  33900
TTTTATCCAT GTCTGTGAAG TCTTTGGACA GAACTGAAAG AGCAACCTCT TTCGGAGGAT  33960
GCCAAAGTGT TGTAGAGTAG ATCTCCATGC CTTCGACTCT GTAATTCTCA ATCCTCCTAA  34020
CCTCTGAGAA TTGTCTTTCA GCTTGCGTGG ACTCTGAAAG TTTACAATAG GCCNTTTCCG  34080
ATTTGGCACA GTACCCAACC GGTATTGCAG TGGTGAGAAG CTAGATGGCT CAAGATGCTG  34140
ATAGCTTCTT TGCCGTGGTA AGAACACAAA GCTAAATAAC CTTTCCCCCT TTCACGAAGA  34200
AGGCTCATCA AGCCTTCCGC TGCTGCTTTT TGTAGATTAA AAGCCTGAAT CTGAGGCGCG  34260
ATTGCGGCTA TTTTCCCTTC TGAAATGACG GAAGAGTCCA ATTTTGTCAC TTCCAGGCTA  34320
TCACTTATGT TCGGTGGAGT TATTGCTCCT TTATTAGTTT TACTTTTGGT TCTTCTGTTT  34380
GGGATTTTAG GTGGAAACTT CATTTTTAAT TTTCTCCTAA TTCTCCTCGG TTGTGGAGCT  34440
GTCACTAGTC AAGAGTCGTG AATTTCTTCG AGGNCGGTGC ATTTGGGGGA GATGCCATAG  34500
TGGGGCTCAA TACCTGAGGT GTTGCCCTTG TCGGCGGACC AGAACTTTGT GTTTTTGCAA  34560
GGACTGGAGT TACCTTTCGG CTCTTTCCCC TCTGCGAGAA GACAGACGGT GTTCCGGTTT  34620
```

```
GGCCGATTCT GGCAACAGGC TTTTCTGAAG GGGCTCCGGT GGATGGCACG TCAGTGACAG  34680
ACGGTGTCTC ATACCAGTGC AGTTTTGTCA ATAGGGTCCG TCTCCGGGAC TTGGGGTTTC  34740
TAATGGCAAA ATGCCAACAC TTGGGGTTAA TGGACTAACA GCTGCTGGTC CTCCTAATAA  34800
ACTTCGACCA GTTTTTGGTT TATGTTGAAC CTGTTTAGAT CATATGGAAG TTCCTGTTCC  34860
CAGTGGGACA GTATCAGGTG AAAGGACAGC TGAATCGATA GAAGACACTG GGGAGTCTGT  34920
ATTCAAGGAG TACTTTGAAT TGGAAGATTC TAAATTCCAT CCGTTTCATT CGACGGTGTC  34980
CTGGGGTGTT TCCGTAAGAA CGGTCTCGGG CTGTCTGTGA CATAAACTAG GACGAGGTCC  35040
AAGTGTTGTG GCGCAACACT TGGACAGGCA GTTGCTAAAG CTCTCTAGAG AGGTGAATCA  35100
AAATGTTTGG TCAGGATCTG GCTTTTCCCC CCTATTTCAC ATCATGATTC AAAGGACAC   35160
CAGAGGAAAG GATTTCAACG AAGGCTCTTT TGGTCACATT CTGATCCTTT GGTAAGCCGA  35220
TCTGTCTTGC AATATACATG TCCCGACGAT GGAAGGGGAA AGCGAGCTGA ATCACCAAAC  35280
TCAGGAACGA TAATATCATC GTGGCTTTTC TGCTTATGAA ACACTCCACC CGATAAGATT  35340
TGATCCCCTT CTGCAAGCTT GCTGAGATCA ACACAACATT TCGCAAGCAG GCATTTGCAT  35400
TGCGGGGTAG TACAACTGTG TCCTTTCAAG AGTCTATATG TTTTATAGGC CTTTCCTGAG  35460
CGGTAAGAAC AGGTCGCCAG TAAGAACAAG GCTTCTTCTG AGTGTACTTC TGCATAAAGG  35520
CGTTCTGCGG GGGAAACCGC ATCTCGGTAG GCATAGTGGT TTAGTGCTTG CCATATAGCA  35580
GCCTGGACGG GTCCCTGCAG CACCGCCATC CTCGAGGCTC AGGCCCACTT TCTGCAGTGC  35640
CACAGGCACC CCCCCCCCCC CATAGCGGCT CCGGCCCGGC CAGCCCCGGC TCATTTAAAG  35700
GCACCAGCCG CCGTTACCGG GGGATGGGGG AGTCCGAGAC AGAATGACTT CTTTATCCTG  35760
CTGACTCTGG AAAGCCCGGC GCCTTGTGAT CCATTGCAAA CCGAGAGTCA CCTCGTGTTT  35820
AGAACACGGA TCCACTCCCA AGTTCAGTGG GGGGATGTGA GGGGTGTGGC AGGTAGGACG  35880
AAGGACTCTC TTCCTTCTGA TTCGGTCTGC ACAGTGGGGC CTAGGGCTGG AGCTCTCTCC  35940
GTGCGGACCG CTGACTCCCT CTACCTTGGG TTCCCTCGGC CCCACCCTGG AACGCCGGGC  36000
CTTGGCAGAT TCTGGCCCTT TCTGGCCCTT CAGTCGCTGT CAGAAACCCC ATCTCATGCT  36060
CGGATGCCCC GAGTGACTGT GGCTCGCACC TCTCCGGAAA CATTGGAAAT CTCTCCTCTA  36120
CGCGCGGCCA CCTGAAACCA CAGGAGCTCG GGACACACGT GCTTTCGGGA GAGAATGCTG  36180
AGAGTCTCTC GCCGACTCTC TCTTGACTTG AGTTCTTCGT GGGTGCGTGG TTAAGACGTA  36240
GTGAGACCAG ATGTATTAAC TCAGGCCGGG TGCTGGTGGC TCACGCCTGT AACCCCAACA  36300
CTTTGGGAGG CCGAGGCCGT AGGATCCCTC GAGGAATCGC CTAACCCTGG GGAGGTTGAG  36360
GTTGCAGTGA GTGAGCCATA GTTGTGTCAC TGTGCTCCAG TCTGGGCGAA AGACAGAATG  36420
AGGCCCTGCC ACAGGCAGGC AGGCAGGCAG GCAGGCAGAA AGACAACAGC TGTATTATGT  36480
TCTTCTCAGG GTAGGAAGCA AAAATAACAG AATACAGCAC TTAATTAATT TTTTTTTTTT  36540
CCTTCGGACG GAGTTTCACT CTTGGTGCCC ACGCTGGAGT GCAGTGGCAC CATCTCGGCT  36600
CACCGCAACC TCCACCTCCC GCGTTCAAGC GATTCTCCTG CCTCAGCCTC CTGAGTAGCT  36660
GGGATTACAG GGAGGAGCCA CCACACCCAG CTGATTTTGT ATTGTTAGTA GAGACGGCAT  36720
TTCTCCATGT GGGTCAGGCT GGTCTCGAAC TGGCGACCCC AGTGGATCTG CCCGCCCCGG  36780
CCTCCCAAAG TGCTGGGGTG ACAGGCGTGA GCCATCGTGA CTGGCCGGCT ACGTTTATTT  36840
ATTTATTTTT TTAATTATTT TACTTTTTTT TAGTTTTCCA TTTTAATCTA TTTATTTATT  36900
TACATTTATT TATTTATTTA TTTATTTACT TATTTATTTA TTTTCGAGAC AGACTCTCGC  36960
TCTGCTGCCC AGGCTGGAGT GCAGCGGCGT GATCTCGGCT CACTGCAACG TCCGCCTCCC  37020
GGGTTCACGC CATTCTCCTG CCTCAGCCTC CCAAGTAGCT GGGACTACAG GCGCCCGCCA  37080
CCGTGCCCGG CTAACTTTTT GTATTTTGAG TAGAGATGGG GTTTCACTGT GGTAGCCAGG  37140
ATGGTCTCGA TCTCCTGACC CCGTGATCCG TCCACCTCGG CCTCCCAAAG TGCTGGGATG  37200
ACAGGCGTGA GCCACCGGCC CCGGCCTATT TATCTATTTA TTAACTTTGA GTCCAGGTTA  37260
TGAAACCAGT TAGTTTTTGT AATTTTTTTT TTTTTTTTTT TTTTTTGAGA CGAGGTTTCA  37320
CCGTGTTGCC AAGGCTTGGA CCGAGGGATC CACCGGCCCT CGGCCTCCCA AAAGTGCGGG  37380
GATGACAGGC GCGAGCCTAC CGCGCCCGGA CCCCCCCTTT CCCCTTCCCC CGCTTGTCTT  37440
CCCGACAGAC AGTTTCACGG CAGAGCGTTT GGCTGGCGTG CTTAAACTCA TTCTAAATAG  37500
AAATTTGGGA CGTCAGCTTC TGGCCTCACG GACTCTGAGC CGAGGAGTCC CCTGGTCTGT  37560
CTATCACAGG ACCGTACACG TAAGGAGGAG AAAAATCGTA ACGTTCAAAG TCAGTCATTT  37620
TGTGATACAG AAATACACGG ATTCACCCAA AACACAGAAA CCAGTCTTTT AGAAATGGCC  37680
TTAGCCCTGG TGTCCGTGCC AGTGATTCTT TTCGGTTTGG ACCTTGACTG AGAGGATTCC  37740
CAGTCGGTCT CTCGTCTCTG GACGGAAGTT CCAGATGATC CGATGGGTGG GGGACTTAGG  37800
CTGCGTCCCC CCAGGAGCCC TGGTCGATTA GTTGTGGGGA TCGCCTTGGA GGGCGCGGTG  37860
ACCCACTGTG CTGTGGGAGC CTCCATCCTT CCCCCCACCC CCTCCCCAGG GGGATCCCAA  37920
TTCATTCCGG GCTGACACGC TCACTGGCAG GCGTCGGGCA TCACCTAGCG GTCACTGTTA  37980
CTCTGAAAAC GGAGGCCTCA CAGAGGAAGG GAGCACCAGG CCGCCTGCGC ACAGCCTGGG  38040
GCAACTGTGT CTTCTCCACC GCCCCCGCCC CCACCTCCAA GTTCCTCCCT CCCTTGTTGC  38100
CTAGGAAATC GCCACTTTGA CGACCGGGTC TGATTGACCT TTGATCAGGC AAAAACGAAC  38160
AAACAGATAA ATAAATAAAA TAACACAAAA GTAACTAACT AAATAAAATA AGTCAATACA  38220
ACCCATTACA ATACAATAAG ATACGATACG ATAGGATGCG ATAGGATACG ATAGGATACA  38280
ATACAATAGG ATACGATACA ATACAATACA ATACAATACA ATACAATACA ATACAATACA  38340
ATACAATACA ATACAATACG CCGGGCGCGG TGGCTCATGC CTGTCATCCC GTCACTTTGG  38400
GATGCCGAGG TGGACGCATC ACCTGAAGTC GGGAGTTGGA GACAAGCCCG ACCAACATGG  38460
AGAAATCCCG TCTCAATTGA AAATACAAAA CTAGCCGGGC GCGGTGGCAC ATGCCTATAA  38520
```

```
TCCCAGCTGC TAGGAAGGCT GAGGCAGGAG AATCGCTTGA ACCTGGGAAG CGGAGGTTGC  38580
AGTGAGCCGA GATTGCGCCA TCGCACTCCA GTCTGAGCAA CAAGAGCGAA ACTCCGTCTC  38640
AAAAATAAAT ACATAAATAA ATACATACAT ACATACATAC ATACATACAT ACATACATAC  38700
ATAAATTAAA ATAAATAAAT AAAATAAAAT AAATAAATGG GCCCTGCGCG GTGGCTCAAG  38760
CCTGTCATCC CCTCACTTTG GGAGGCCAAG GCCGGTGGAT CAAGAGGCGG TCAGACCAAC  38820
AGGGCCAGTA TGGTGAAACC CCGTCTCTAC TCACAATACA CAACATTAGC CGGGCGCTGT  38880
GCTGTGCTGT ACTGTCTGTA ATCCCAGCTA CTCGGGAGGC CGAGCTGAGG CAGGAGAATC  38940
GCTTGAACCT GGGAGGCGGA GGTTGCAGTG AGCCGAGATC GCGCCACTGC AACCCAGCCT  39000
GGGCGACAGA GCGAGACTCC GTCTCCAAAA AATGAAAATG AAAATGAAAC GCAACAAAAT  39060
AATTAAAAAG TGAGTTTCTG GGGAAAAAGA AGAAAAGAAA AAAGAAAAAA ACAACAAAAC  39120
AGAACAACCC CACCGTGACA TACACGTACG CTTCTCGCCT TTCGAGGCCT CAAACACGTT  39180
AGGAATTATG CGTGATTTCT TTTTTTAACT TCATTTTATG TTATTATCAT GATTGATGTT  39240
TCGAGACGGA GTCTCGGAGG CCCGCCCTCC CTGGTTGCCC AGACAACCCC GGGAGACAGA  39300
CCCTGGCTGG GCCCGATTGT TCTTCTCCTT GGTCAGGGGT TTCCTTGTCT TTCTTCGTGT  39360
CTTTAACCCG CGTGGACTCT TCCGCCTCGG GTTTGACAGA TGGCAGCTCC ACTTTAGGCC  39420
TTGTTGTTGT TGGGGACTTT CCTGATTCTC CCCAGATGTA GTGAAAGCAG GTAGATTGCC  39480
TTGCCTGGCC TTGCCTGGCC TTGCCTTTTC TTTCTTTCTT TCTTTCTTTA TTACTTTCTC  39540
TTTTTCTTCT TCTTCTTCTT CTTTTTTTTG AGACAGAGTT TCACTCTTGT TGCCCAGGCT  39600
AGAGGGCAAT GGCGCGATCT CGGCTCACCG CACCCTCCGC CTCCCAGGTT CAAGCGATTC  39660
TCCTGCCTCA GCCTCCTGAT TAGCTGGGAT TACAGGCATG GGCCACCGTG CTGGCTGATG  39720
TTTGTACTTT TAGTAGAGAC GGTGTTTTTC CATGTTGGTC AGGCTGGTCT CCCACTCCCA  39780
ACCTCAGGTG GTCCGCCTGC CTTAGCCTCC CAAAGTGCTG GGATGACAGG CGTGCAACCG  39840
CGCCCAGCCT CTCTCTCTCT CTCTCTCTCT CTCGCTCGCT TGCTTGCTTG CTTTCGTGCT  39900
TTCTTGCTTT CCCGTTTTCT TGCTTTCTTT CTTTCTTTCG TTTCTTTCAT GCTTGCTTTC  39960
TTGCTTGCTT GCTTGCTTTC GTGCTTTCTT GCTTTCCTGT TTTCTTTCTT TCTTTCTTTC  40020
TTTCTTTCTT TTGTTTCTTT CTTGCTTGCT TTCTTGCTTG CTTGCTTGCT TTCGTGCTTT  40080
CTTGCTTTCC TGTTTTCTTT CTTTCTTTCT TTCTTTTCTT TCTTTCTTGC TTGCTTTCCT  40140
GCTTGCTTGC TTTCGTGCTT TCTTGTTTTC TCGATTTCTT TCTTTCTTTT GTTTCTTTCC  40200
TGCTTGCTTT CTTGCTTGCT TGCTTTCGTG CTTCTTGCTT TCCTGTTTTC TTTCTTTCTT  40260
TCTTTCTTTT GTTTCTTTCT TGCTTGCTTT CTTGCTTGCT TGCTTTCGTG CTGTCTTGTT  40320
TCTCGATTTC TTTCTTTCTT TTGTTTCTTT CCTGCTTGCT TTCTTGCTTG ATTGCTTTCG  40380
TGCTTTCTTG CTTTCTTGTT TTCTTTCTTT CTTTTGTTTC TTTCTTTCTT GCTTCCTTGT  40440
TTTCTTGCTT TCTTGCTTGC TTGCTTTCGT GCTTTCTTGT TTTCTTGCTT TCTTTCTTTT  40500
GTTTCTTTCT TGCTTGCTTT CTTGCTTCCT TGTTTTCTTG CTTTCTTGCT TGCTTGCTTT  40560
CGTGCTTTCT TTCTTGCTTT CTTTTCTTTC TTTCTTTTCT TTTTCTTTCT TTCTTGCTTT  40620
CTTTTCTTTC ATCATCATCT TTCTTTCTTT CCTTTCTTTC TTTCTTTCTT TCTATCTTTC  40680
TTTCTTTCTT TCTTTCTTTC TTTCTTTCTT TCTTTCTGTT TCGTCCTTTT GAGACAGAGT  40740
TTCACTCTTG TTTCCACGGC TAGAGTGCAA TGGCGCGATC TTGGCTCACC GCACCTTCCG  40800
CCTCCCGGGT TCGAGCGCTT CTCCTGCCTC CAGCCTCCCG ATTAGCGGGG ATTGACAGGG  40860
AGGCACCCCC ACGCCTGGCT TGGCTGATGT TTGTGTTTTT AGTAGGCACG CCGTGTCTCT  40920
CCATGTTGCT CAGGCTGGTC TCCAACTCCC GACCTCCTGT GATGCGCCCA CCTCGGCCTC  40980
TCGAAGTGCT GGGATGACGG GCGTGACGAC CGTGCCCGGC CTGTTGACTC ATTTCGCTTT  41040
TTTATTTCTT TCGTTTCCAC GCGTTTACTT ATATGTATTA ATGTAAACGT TTCTGTACGC  41100
TTATATGCAA ACAACGACAA CGTGTATCTC TGCATTGAAT ACTCTTGCGT ATGGTAAATA  41160
CGTATCGGTT GTATGGAAAT AGACTTCTGT ATGATAGATG TAGGTGTCTG TGTTATACAA  41220
ATAAATACAC ATCGCTCTAT AAAGAAGGGA TCGTCGATAA AGACGTTTAT TTTACGTATG  41280
AAAAGCGTCG TATTTATGTG TGTAAATGAA CCGAGCGTAC GTAGTTATCT CTGTTTTCTT  41340
TCTTCCTCTC CTTCGTGTTT TTCTTCCTTC CTTTCTTCCT TTCTCTCCTT CTTTAGGTTT  41400
TTCTTCCTCT CTTCCTTTCC TTCTTTCTCT CTTTCTGTCC TTTTTTCCTT CGTGCTTTAT  41460
TTCTCTTTCG TTCCCTGTGT TTCCTTCTTT TTTCTTTCCT CTCTGTTTCT TTTTCCCTTC  41520
TTTCCTTCGT TTCTTTCCTC ATTCTTTCTC TCTTTTTCGT TGTTTCTTTC CTTCCCGTCT  41580
GTCTTTTAAA AAATTGGAGT GTTTCAGAAG TTTACTTTGT GTATCTACGT TTTCTAAATT  41640
GTCTCTCTTT TCTCCATTTT CTTCCTCCCT CCCTCCCTCC CTCCCTGCTC CCTTCCCTCC  41700
CTCCTTCCCT TTCGCCATCT GTCTCTTTTC CCCACTCCCC TCCCCCCGTC TGTCTCTGCG  41760
TGGATTCCGG AAGAGCCTAC CGATTCTGCC TCTCCGTGTG TCTGCAGCGA CCCCGCGACC  41820
GAGTCCTTGT GTGTTCTTTC TCCCTCCCTC CCTCCCTCCC TCCCTCCCTC CCTCCCTGCT  41880
TCCGAGAGGC ATCTCCAGAG ACCGCGCCGT GGGTTGTCTT CTGACTCTGT CGCGGTCGAG  41940
GCAGAGACGC GTTTTGGGCA CCGTTTGTGT GGGGTTGGGG CAGAGGGGCT GCGTTTTCGG  42000
CCTCGGGAAG AGCTTCTCGA CTCACGGTTT CGCTTTCGCG GTCCACGGGC CGCCCTGCCA  42060
GCCGGATCTG TCTCGCTGAC GTCCGCGGCG GTTGTCGGGC TCCATCTGGC GGCCGCTTTG  42120
AGATCGTGCT CTCGGCTTCC GGAGCTGCGG TGGCAGCTGC CGAGGGAGGG GACCGTCCCC  42180
GCTGTGAGCT AGGCAGAGCT CCGGAAAGCC CGCGGTCGTC AGCCCGGCTG GCCCGGTGGC  42240
GCCAGAGCTG TGGCCGGTCG CTTGTGAGTC ACAGCTCTGG CGTGCAGGTT TATGTGGGGG  42300
AGAGGCTGTC GCTGCGCTTC TGGGCCCGCG GCGGGCGTGG GGCTGCCCGG GCCGGTCGAC  42360
CAGCGCGCCG TAGCTCCCGA GGCCCGAGCC GCGACCCGGC GGACCCGCCG CGCGTGGCGG  42420
```

```
AGGCTGGGGA CGCCCTTCCC GGCCCGGTCG CGGTCCGCTC ATCCTGGCCG TCTGAGGCGG  42480
CGGCCGAATT CGTTTCCGAG ATCCCCGTGG GGAGCCGGGG ACCGTCCCGC CCCCGTCCCC  42540
CGGGTGCCGG GGAGCGGTCC CCGGGCCGGG CCGCGGTCCC TCTGCCGCGA TCCTTTCTGG  42600
CGAGTCCCCG TGGCCAGTCG GAGAGCGCTC CCTGAGCCGG TGCGGCCCGA GAGGTCGCGC  42660
TGGCCGGCCT TCGGTCCCTC GTGTGTCCCG GTCGTAGGAG GGGCCGGCCG AAAATGCTTC  42720
CGGCTCCCGC TCTGGAGACA CGGGCCGGCC CCTGCGTGTG GCCAGGGCGG CCGGGAGGGC  42780
TCCCCGGCCC GGCGCTGTCC CCGCGTGTGT CCTTGGGTTG ACCAGAGGGA CCCCGGGCGC  42840
TCCGTGTGTG GCTGCGATGG TGGCGTTTTT GGGGACAGGT GTCCGTGTCC GTGTCGCGCG  42900
TCGCCTGGGC CGGCGGCGTG GTCGGTGACG CGACCTCCCG GCCCCGGGGG AGGTATATCT  42960
TTCGCTCCGA GTCGGCAATT TTGGGCCGCC GGGTTATAT                         42999
```

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 175 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

```
CTCCCGCGCG GCCCCCGTGT TCGCCGTTCC CGTGGCGCGG ACAATGCGGT TGTGCGTCCA     60
CGTGTGCGTG TCCGTGCAGT GCCGTTGTGG AGTGCCTCGC TCTCCTCCTC CTCCCCGGCA    120
GCGTTCCCAC GGTTGGGGAC CACCGGTGAC CTCGCCCTCT TCGGGCCTGG ATCCG         175
```

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 755 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

```
GGTCTGGTGG GAATTGTTGA CCTCGCTCTC GGGTGCGGCC TTTGGGGAAC GGCGGGGTCG     60
GTCGTGCCCG GCGCCGGACG TGTGTCGGGG CCCACTTCCC GCTCGAGGGT GGCGGTGGCG    120
GCGGCGTTGG TAGTCTCCCG TGTTGCGTCT TCCCGGGCTC TTGGGGGGGG TGCCGTCGTT    180
TTCGGGGCCG GCGTTGCTTG GCTTACGCAG GCTTGGTTTG GGACTGCCTC AGGAGTCGTG    240
GGCGGTGTGA TTCCCGCCGG TTTTGCCTCG CGTCTGCCTG CTTTGCCTCG GGTTTGCTTG    300
GTTCGTGTCT CGGGAGCGGT GGTTTTTTTT TTTTTCGGGT CCCGGGGAGA GGGGTTTTTC    360
CGGGGGACGT TCCCGTCGCC CCCTGCCGCC GGTGGGTTTT CGTTTCGGGC TGTGTTCGTT    420
TCCCCTTCCC CGTTTCGCCG TCGGTTCTCC CCGGTCGGTC GGCCCTCTCC CCGGTCGGTC    480
GCCCGGCCGT GCTGCCGGAC CCCCCCTTCT GGGGGGGATG CCCGGGCACG CACGCGTCCG    540
GGCGGCCACT GTGGTCCGGG AGCTGCTCGG CAGGCGGGTG AGCCAGTTGG AGGGGCGTCA    600
TGCCCCCGCG GGCTCCCGTG GCCGACGCGG CGTGTTCTTT GGGGGGGCCT GTGCGTGCGG    660
GAAGGCTGCG CACGTTGTCG GTCCTTGCGA GGGAAAGAGG CTTTTTTTTT TTAGGGGGTC    720
GTCCTTCGTC GTCCCGTCGG CGGTGGATCC GGCCT                               755
```

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 463 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

```
GGCCGAGGTG CGTCTGCGGG TTGGGGCTCG TCCGGCCCCG TCGTCCTCCG GGAAGGCGTT     60
TAGCGGGTAC CGTCGCCGCG CCGAGGTGGG CGCACGTCGG TGAGATAACC CCGAGCGTGT    120
TTCTGGTTGT TGGCGGCGGG GGCTCCGGTC GATGTCTTCC CCTCCCCCTC TCCCCGAGGC    180
CAGGTCAGCC TCCGCCTGTG GGCTTCGTCG GCCGTCTCCC CCCCCCTCAC GTCCCTCGCG    240
AGCGAGCCCG TCCGTTCGAC CTTCCTTCCG CCTTCCCCCC ATCTTTCCGC GCTCCGTTGG    300
CCCCGGGGTT TTCACGGCGC CCCCCACGCT CCTCCGCCTC TCCGCCCGTG GTTTGGACGC    360
CTGGTTCCGG TCTCCCCGCC AAACCCCGGT TGGGTTGGTC TCCGGCCCCG GCTTGCTCTT    420
CGGGTCTCCC AACCCCCGGC CGGAAGGGTT CGGGGGTTCC GGG                      463
```

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 378 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

```
GGATTCTTCA GGATTGAAAC CCAAACCGGT TCAGTTTCCT TTCCGGCTCC GGCCGGGGGG     60
GGCGGCCCCG GGCGGTTTGG TGAGTTAGAT AACCTCGGGC CGATCGCACG CCCCCCGTGG    120
CGGCGACGAC CCATTCGAAC GTCTGCCCTA TCAACTTTCG ATGGTAGTCG ATGTGCCTAC    180
CATGGTGACC ACGGGTGACG GGGAATCAGG GTTCGATTCC GGAGAGGGAG CCTGAGAAAC    240
GGCTACCACA TCCAAGGAAG GCAGCAGGCG CGCAAATTAC CCACTCCCGA CCCGGGGAGG    300
TAGTGACGAA AAATAACAAT ACAGGACTCT TTCGAGGCCC TGTAATTGGA ATGAGTCCAC    360
TTTAAATCCT TTAAGCAG                                                  378
```

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 378 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

```
GATCCATTGG AGGGCAAGTC TGGTGCCAGC AGCCGCGGTA ATTCCAGCTC CAATAGCGTA     60
TATTAAAGTT GCTGCAGTTA AAAAGCTCGT AGTTGGATCT TGGGAGCGGG CGGGCGGTCC    120
GCCGCGAGGC GAGTCACCGC CCGTCCCCGC CCCTTGCCTC TCGGCGCCCC CTCGATGCTC    180
```

```
TTAGCTGAGT TGTCCCGCGG GGCCCGAAGC GTTTACTTTG AAAAAATTAG AGTTGTTTCA     240
AAGCAGGCCC GAGCCGCCTG GATACCGCCA GCTAGGAAAT AATGGAATAG GACCGCGGTT     300
CCTATTTTGT TTGGTTTTCG GAACTGAGCC CATGATTAAG GGAAACGGCC GGGGGCATTC     360
CCTTATTGCG CCCCCCTA                                                   378
```

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 719 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

```
GGATCTTTCC CGCTCCCCGT TCCTCCCGGC CCCTCCACCC GCGCGTCTCC CCCCTTCTTT      60
TCCCCTCTCC GGAGGGGGGG GAGGTGGGGG CGCGTGGGCG GGGTCGGGGG TGGGGTCGGC     120
GGGGGACCGC CCCCGGCCGG CAAAAGGCCG CCGCCGGGCG CACTTCAACC GTAGCGGTGC     180
GCCGCGACCG GCTACGAGAC GGCTGGGAAG GCCCGACGGG GAATGTGGCT CGGGGGGGGC     240
GGCGCGTCTC AGGGCGCGCC GAACCACCTC ACCCCGAGTG TTACAGCCCT CCGGCCGCGC     300
TTTCGCGGAA TCCCGGGGCC GAGGGGAAGC CCGATACCCG TCGCCGCGCT TTTCCCCTCC     360
CCCCGTCCGC CTCCCGGGCG GGCGTGGGGG TGGGGGCCGG GCCGCCCCTC CCACGCCCGT     420
GGTTTCTCTC TCTCCCGGTC TCGGCCGGTT TGGGGGGGGG AGCCCGGTTG GGGGCGGGGC     480
GGACTGTCCT CAGTGCGCCC CGGGCGTCGT CGCGCCGTCG GGCCCGGGGG GTTCTCTCGG     540
TCACGCCGCC CCCGACGAAG CCGAGCGCAC GGGGTCGGCG GCGATGTCGG CTACCCACCC     600
GACCCGTCTT GAAACACGGA CCAAGGAGTC TAACGCGTGC GCGAGTCAGG GGCTCGCACG     660
AAAGCCGCCG TGGCGCAATG AAGGTGAAGG GCCCCGTCCG GGGGCCCGAG GTGGGATCC      719
```

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 685 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

```
CGAGGCCTCT CCAGTCCGCC GAGGGCGCAC CACCGGCCCG TCTCGCCCGC CGCGTCGGGG      60
AGGTGGAGCA CGAGCGTACG CGTTAGGACC CGAAAGATGG TGAACTATGC CTGGGCAGGG     120
CGAAGCCAGA GGAAACTCTG GTGGAGGTCC GTAGCGGTCC TGACGTGCAA ATCGGTCGTC     180
CGACCTGGGT ATAGGGGCGA AAGACTAATC GAACCATCTA GTAGCTGGTT CCCTCCGAAG     240
TTTCCCTCAG GATAGCTGGC GCTCTCGCAA CCTTCGGAAG CAGTTTTATC CGGGTAAAGG     300
CGGAATGGAT TAGGAGGTCT TGGGGCCGGA AACGATCTCA AACTATTTCT CAAACTTTAA     360
ATGGGTAAGG AAGCCCGGCT CGCTGGCGTG GAGCCGGGCG TGGAATGCGA GTGCCTAGTG     420
GGCCACTTTT GGTAAGCAGA ACTGGCGCTG CGGGATGAAC CGAACGCCGG GTTAAGGCGC     480
CCGATGCCGA CGCTCATCAG ACCCCAGAAA AGGTGTTGGT TGATATAGAC AGCAGGACGG     540
TGGCCATGGA AGTCGGAATC CGCTAAGGAG TGTGTAACAA CTCACCTGCC GAATCAACTA     600
GCCCTGAAAA TGGATGGCGC TGGAGCGTCG GGCCCATACC CGGCCGTCGC CGGCAGTCGG     660
AACGGGACGG GACGGGAGCG GCCGC                                           685
```

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 33 base pairs
  (B) TYPE: nucleic acid
  (C) STRANDEDNESS: single
  (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

GAGGAATTCC CCTATCCCTA ATCCAGATTG GTG 33

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 35 base pairs
  (B) TYPE: nucleic acid
  (C) STRANDEDNESS: single
  (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

AAACTGCAGG CCGAGCCACC TCTCTTCTGT GTTTG 35

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 33 base pairs
  (B) TYPE: nucleic acid
  (C) STRANDEDNESS: single
  (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

AGGAATTCAC AGAAGAGAGG TGGCTCGGCC TGC 33

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 34 base pairs
  (B) TYPE: nucleic acid
  (C) STRANDEDNESS: single
  (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

AGCCTGCAGG AAGTCATACC TGGGGAGGTG GCCC 34

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 80 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

AAACTGCAGG TTAATTAACC CTAACCCTAA CCCTAACCCT AACCCTAACC CTAACCCTAA 60
CCCTAACCCT AACCCGGGAT 80

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

TTGGGCCCTA GGCTTAAGG 19

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

GCCAGGGTTT TCCCAGTCAC GACGT 25

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

GCTGCAAGGC GATTAAGTTG GGTAAC 26

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

TATGTTGTGT GGAATTGTGA GCGGAT 26

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

GGGTTTAAAC AGATCTCTGC A 21